AN ANALYSIS OF OWNER WEALTH MAXIMIZING VERSUS EQUAL OUTCOME RESIDENTIAL MORTGAGE LENDING

By

CHRISTOPHER L. BROWN

Bachelor of Business Administration
University of Central Arkansas
Conway, Arkansas
1983

Master of Business Administration
University of Central Arkansas
Conway, Arkansas
1986

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1998

AN ANALYSIS OF OWNER WEALTH MAXIMIZING

VERSUS EQUAL OUTCOME RESIDENTIAL

MORTGAGE LENDING

Dissertation Approved:

W. Gary Simpson

Dissertation Advisor

Jim Krehbiel

Wayne B Powell

Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my dissertation chair, Dr. Gary Simpson, and dissertation committee members, Dr. John Polonchek, Dr. Tim Krehbiel, and Dr. Rick Wilson. These individuals have given me excellent guidance, assistance, and encouragement on my dissertation and throughout my tenure as a graduate student at Oklahoma State University. I would like to give special thanks to Dr. Gary Simpson for his close supervision and for the time he has devoted to answering all of my questions. I would like to thank Dr. Janice Jadlow and the Department of Finance for providing me with financial assistance during the first four years of my graduate studies at Oklahoma State University.

I would like to thank Dr. Stan Liebowitz for providing additional variables for the loan application dataset. Without the additional variables, the completion of this dissertation would not have been possible.

I would like to give special attention and appreciation to my wife, Laurie, for her constant support and encouragement and to my two sons, William and Jared, who provided a constant reminder of the importance of finishing my graduate work. Finally, I would like to thank God for giving me the strength, perseverence, and wisdom to complete my graduate work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I.
INTRODUCTION

A. The Research Problem

The issue of discrimination in residential mortgage lending has received a great deal of attention. Numerous reports in the popular press have addressed the issue and the evidence consistently suggests denial rates for home loans are higher for minorities than for nonminorities.[1] Public officials have reacted strongly to the appearance of widespread racial discrimination in mortgage lending. The press reports prompted investigations by the Justice Department and the House Banking Committee (Cocheo (1993)). The Clinton administration and banking regulators have made public statements which indicate deep concern about racial discrimination in the extension of credit (Cocheo (1993) and Bacon (1993)). The possibility of widespread racial discrimination in mortgage lending has serious public policy implications for the regulation of the financial system and the management of financial institutions that extend mortgage credit.

Much of the attention can be attributed to the analysis of the data reported as a result of the 1989 and 1991 amendments to the Home Mortgage Disclosure Act (HMDA).[2] The HMDA requires mortgage lenders to compile and report data from home

---

[1] The *Washington Post* reported white neighborhoods receive twice as many mortgage loans as black neighborhoods with equivalent income levels (Brenner and Spayd (1993)). A similar analysis of the Atlanta mortgage market came to the same conclusion (*Atlanta Constitution (1994)*.

[2] An analysis by the Boston Federal Reserve Bank of HMDA data indicated that minorities are denied loans at twice the rate of nonminorities, even when factors such as income are held constant (Munnell, *et. al.* (1996)).

purchase and home improvement loan applications. These disclosures are meant to help authorities identify discriminatory lending practices and enforce the following statutes: (1) the Fair Housing Act, (2) the Equal Credit Opportunity Act (ECOA), and (3) the Community Reinvestment Act (CRA). However, the accurate investigation of possible discrimination in lending with HMDA data is a complex problem because several crucial factors in the credit decision not included in the HMDA data are highly correlated with race.[3]

Recent statements from policymakers in Washington indicate that performance-based standards will be used to assess lenders' performance in extending credit to minorities (England (1993)). President Clinton has indicated that lenders should be evaluated by their performance rather than their effort in lending to minorities (England (1993)). Critics argue that performance-based standards are a move toward lending quotas and equal outcomes for all races, regardless of credit quality (England (1993)).

Equal outcome lending will accomplish the social goal of eliminating racial disparities in lending. The cost of this type of lending program to society depends on whether the racial disparities that exist are due primarily to discrimination or whether they occur because nonminority applicants tend to be more creditworthy than minority applicants.

Disparities in credit extension to minorities and nonminorities could result from three basic causes. First, disparities could result from lenders having a "taste for

[3] Munnell, *et. al.* (1996) find evidence that blacks are more likely to have lower net worths, higher monthly obligations to income ratios, and worse credit histories than nonminorities.

discrimination."[4] This implies lenders deny profitable loans to minority applicants to satisfy their "taste for discrimination." This type of discrimination is inconsistent with the concept of owner wealth maximization and damages both shareholders and customers. It is also inconsistent with recent empirical evidence which indicates that highly qualified minority borrowers are just as likely to receive mortgage money as highly qualified white borrowers (Hunter and Walker (1995)).

Second, disparities in credit extension between minorities and nonminorities may be the result of owner wealth maximizing behavior by lenders. Lenders should approve all positive risk-adjusted NPV loans to maximize shareholder wealth. If nonminorities tend to be more creditworthy than minorities, lending firms pursuing owner wealth maximizing decisions will make fewer loans to minorities than to nonminorities. The lending firm does not reject positive risk-adjusted NPV loans because of prejudice, but the process results in racial disparities in lending decisions.

The case of owner wealth maximizing credit decisions which result in racial disparities in credit extension raises the classic dilemma of credit rationing. Society must decide through the political process how adjustments should be made to redirect credit to preferred sectors of the economy. If privately owned lenders are asked to make loans to borrowers with higher default risk at market rates, the costs will be borne by the bank's owners and/or other customers assuming no government subsidies are given.

[4] Becker (1971) argues that individuals with a "taste for discrimination" willingly forego profitable investment opportunities to avoid conducting business with minorities.

Third, racial disparities in lending decisions could result from statistical discrimination.[5] Statistical discrimination occurs when a lender either uses race as a proxy for creditworthiness or uses variables highly correlated with race that do not accurately reflect credit risk as proxies for creditworthiness.[6] Statistical discrimination results in higher loan denial rates for minority borrowers than are justified based on relevant measures of credit quality.

Statistical discrimination may or may not be intentional. If lenders use race as a proxy for creditworthiness, the discrimination is intentional. Unintentional statistical discrimination can result from a faulty credit decision model. If extraneous variables are included in the credit decision model that do not influence default risk, some positive risk-adjusted NPV loans may be rejected.[7] If these extraneous variables are correlated with race, the distortions caused by the model will impact minorities disproportionately. The result is statistical discrimination and lending that does not maximize owner wealth because positive risk-adjusted NPV loans are rejected.

---

[5] Statistical discrimination is a term that has been used by many researchers to refer to lenders using unsuitable proxies for credit risk that result in higher denial rates for minority borrowers. See for example, Longhofer (1996), and Ferguson and Peters (1997).

[6] This type of discrimination is referred to legally as disparate impact discrimination. Disparate impact discrimination occurs "when a lender applies a practice uniformly to all applicants but the practice is not justified by business necessity," and it has a disproportionate impact on a protected group (Bauer and Cromwell (1994)).

[7] Lenders may believe these extraneous variables influence default risk. Therefore, they feel justified in including the variables. The statistical discrimination that results from the model is unintentional.

Unintentional statistical discrimination resulting from an erroneous credit decision process can be corrected if lenders use only those factors that accurately reflect the probability of default. This will eliminate the problems associated with extraneous economic factors that are highly correlated with race but racial disparities will still exist if economic factors that truly reflect the probability of default are highly correlated with race.

The ideal solution to this problem would be a credit decision model which contains only those variables required to accurately reflect the probability of default but no variables that would produce racial disparities in the loan decisions. The existence of an owner wealth maximizing, equal outcome credit model is dependent on the degree of correlation of race with economic variables that accurately reflect default risk.

If nonminority borrowers have higher average credit quality than minority borrowers, an owner wealth maximizing lending model will have higher loan approval rates for nonminority borrowers than for minority borrowers. If this is the case, performance-based standards encouraging equal outcomes will lead to lending that is not owner wealth maximizing. Policymakers should consider the costs of implementing these standards.

The research problem is to empirically explore the possibility of a realistic credit decision model that selects all positive NPV loans without producing racial disparities in loan decisions. If race is highly correlated with important default variables, then the ideal owner wealth maximizing, equal outcome model will probably not exist. The existence of an owner wealth maximizing, equal outcome model is ultimately an empirical question and is the critical question addressed in this research. Given the possibility that an ideal credit decision model does not really exist, what is the cost of equal outcome lending?

B. *Purpose of the Research*

The purpose of this research is to develop an empirical owner wealth maximizing credit decision model for mortgage loans which can be adjusted to prevent racial disparities in lending decisions. The equal outcome model is compared to the model that accepts all risk-adjusted positive NPV loans, regardless of racial disparities. If both models predict loan defaults equally well, this would be evidence supporting the possibility of an owner wealth maximizing, equal outcome mortgage lending model. If a difference is detected in default prediction, the costs and benefits of the two models can be compared.

C. *Structure of the Research*

First, a theoretical model of the credit extension decision that accepts all positive risk-adjusted NPV loans is developed. This theoretical model does not address the issue of racial disparities. Only those factors that theory indicates should affect loan approvals are included, regardless of their possible correlation with race. Next, the theoretical model is estimated empirically with a set of loan approval data to test the performance of the owner wealth maximizing model in reducing racial disparities in loan approvals. A second empirical model is developed by testing all possible subsets of the owner wealth maximizing model to find a model that eliminates racial disparities in loan approvals.

Finally, the two models are estimated with a new data set that is based on loan defaults. The cost of using the equal outcome model is any decline in default prediction rates from using the equal outcome model instead of the owner wealth maximizing model and the benefits are considered to be any decline in the racial disparity in loan approvals from using the equal outcome model instead of the owner wealth maximizing model.

This research does not try to measure the social benefits of a political economic policy that requires equal lending across all racial groups. Such a policy would certainly provide benefits to some segments of society, but the benefits to society as a whole would depend on how equal lending across racial groups is achieved.

The dissertation is organized as follows: Chapter II is a review of the previous research on discrimination in mortgage lending, Chapter III reviews the previous research on loan default models, Chapter IV develops a theoretical framework for credit decisions that maximize owner wealth, Chapter V describes the empirical analysis and the hypotheses, Chapter VI presents the results of the empirical analysis, and Chapter VII develops the final conclusions and implications of the research.

CHAPTER II.
THE EVIDENCE ON DISCRIMINATION IN MORTGAGE LENDING

This chapter reviews the evidence on discrimination in mortgage lending. It is well known that there are racial disparities in residential mortgage loan originations and that upper-income neighborhoods receive significantly more residential mortgage loans (measured both in the number of loans and the dollar amount of loans) than lower-income neighborhoods. Canner and Passmore (1995) review the 1993 Home Mortgage Disclosure Act (HMDA) data and report that 90 percent of the total home purchase loans and 92 percent of the total dollar amount of home purchase loans are made in middle- and upper-income neighborhoods. They find the ratio of home purchase loans to the number of owner occupied housing units is significantly lower in low-income neighborhoods than in upper-income neighborhoods. Canner and Passmore (1995) also report rejection rates of 22.25 percent for black applicants applying for government-insured loans and 34.02 percent for black applicants applying for conventional loans. The rejection rates for white applicants are reported to be 11.78 percent for government-insured loans and 15.33 percent for conventional loans. Canner, Passmore and Smith (1994) review the 1992 HMDA data and report similar results. The data suggest that black applicants have significantly higher rejection rates than white applicants and that low-income neighborhoods receive significantly less mortgage money than upper-income neighborhoods. Are these disparities the result of discrimination, differences in the quality of applicants based on objective economic criteria, or both? That is the question that credit discrimination research attempts to answer.

A. *Theories of Credit Rationing and Discrimination in Mortgage Lending*

Credit rationing occurs when lenders charge a bank-optimal interest rate that is below the market-clearing interest rate (Stiglitz and Weiss (1981)). When lenders charge the same rate to all borrowers and that rate is below the market-clearing level, some applicants will be rationed out of the credit market. Lenders must be able to differentiate low-risk applicants from high-risk applicants to ensure that high-risk applicants will be rationed out of the market and low-risk applicants will obtain loans.

Nesiba (1996) provides an example of credit rationing when there are two distinguishable groups of borrowers. Nesiba (1996) assumes the two groups of borrowers are white borrowers and black borrowers. Assume the bank sets a bank-optimal interest rate below the market clearing level so that some borrowers will be rationed out of the mortgage market. Further assume that the projects of black and white borrowers have equal expected returns but the projects of black borrowers have a higher average variance. If the bank-optimal interest rate is low enough, Nesiba (1996) argues that no black borrowers will be given loans until most (or all) of the white borrowers are given loans. Nesiba (1996) theorizes that the combination of imperfect information and two distinct groups with different average risk may lead to rational discrimination against the group with higher average risk. In the context of credit rationing, this is not discrimination. Black borrowers are simply rationed out of the market because they are, on average, higher credit risks.

Credit rationing of the type just described may be rational, but it is also illegal. This type of discrimination has come to be known as rational or statistical discrimination.

Statistical discrimination occurs when a lender uses an attribute, such as race, as a proxy for creditworthiness. The true creditworthiness of the borrower is not influenced by race, but race is correlated with unobservable factors of creditworthiness. Therefore, lenders may use race as an inexpensive proxy for the unobservable factors of creditworthiness.

Becker (1971) argues that some individuals have a taste for discrimination. These individuals are willing to pay or forfeit income to keep from associating with certain groups of people. Becker (1971) develops a model of the labor market with discrimination of this type.

Nesiba (1996) extends the Becker model of "taste-based" discrimination to the mortgage market. Assume lenders make loans with expected repayment R, based on the principal amount of the loan, P, and the interest rate charged on the loan, i:

$$R = P(1 + i).$$

Nesiba (1996) states that discrimination would result in minority borrowers being charged a higher interest rate, $i^* > i$. The difference between the rate charged to minority borrowers and the rate charged to white borrowers $(i^* - i)$ is the discrimination coefficient. The cost of discrimination is $P(i^* - i)$ and is paid by the person being discriminated against. The cost to the discriminating lender would be the income lost from minority applicants that drop out of the applicant pool as a result of being charged a higher interest rate.

As Nesiba (1996) points out, the Becker theory of discrimination assumes a perfectly competitive market. In this perfectly competitive market, if lenders discriminate against minorities by charging higher interest rates, nondiscriminating competitors will enter the market and make loans to minorities at competitive interest rates. The effect is that any discrimination against minorities must be temporary. It cannot continue in the long-term in a perfectly competitive market.

Becker (1993) argues that, if lenders have a taste for discrimination, minority borrowers will be held to a higher standard than white borrowers. The result should be that minorities will have lower default rates than white borrowers. The empirical evidence suggests that black borrowers have higher default rates than white borrowers. Becker argues that the empirical evidence is inconsistent with discrimination against black borrowers.

Notice, this is different from the argument made by Nesiba (1996). If the discrimination takes the form described by Nesiba (1996), the higher interest rate charged to minorities would increase adverse selection and moral hazard problems.[8] The result of this type of discrimination could very well be higher default rates for black borrowers, since high quality black applicants will drop out of the applicant pool.

If discrimination takes the form described by Becker (1993), the result will be lower default rates for minorities only if certain conditions hold. Default rates will only be lower for minorities under this form of discrimination if white borrowers and minority borrowers have identical distributions of credit quality. If white borrowers have higher distributions of credit quality than minority borrowers, minorities will have higher default

[8] It is not necessary for the minority borrowers to be charged a higher interest rate to increase problems of adverse selection and moral hazard. These problems will also exist if other costs of the loan increase, such as mortgage insurance premiums.

rates than white borrowers even if they are held to a higher standard in the loan approval process.[9]

Calem and Stutzer (1995) develop an equilibrium model of credit rationing where minority applicants will have higher denial rates than white applicants and minority borrowers will have higher default rates than white borrowers. The authors assume that applicants of very high credit quality and very low credit quality are easily distinguished by lenders. Problems occur in evaluating applicants with intermediate (marginal) credit scores. Lenders are unable to clearly distinguish high risk applicants from low risk applicants within the intermediate credit scores.

Calem and Stutzer (1995) assume high risk applicants have more to lose if their loan is denied due to their higher probability of default. Therefore, higher risk applicants are more willing to accept more expensive loans if it improves the chance of loan approval. This is consistent with previous models of credit rationing, such as Stiglitz and Weiss (1981), where charging higher rates leads to a lower quality applicant pool.

Calem and Stutzer (1995) assume some lenders offer higher cost loans while other lenders offer lower cost loans.[10] The authors also assume that lenders offering higher cost

---

[9] This relationship between default and denial rates is developed formally by Ferguson and Peters (1995). Tootell (1993), Browne and Tootell (1995), and Galster (1993) also make this point in responding to criticisms by Becker (1993).

[10] High cost loans may be FHA loans where the borrower must pay a mortgage insurance premium, while low cost loans may be conventional loans with loan-to-value ratios below 80 percent. If the required downpayment is not considered a cost of the loan, FHA loans are significantly more costly than conventional loans with low loan-to-value ratios. Empirical evidence suggests that minority borrowers tend to be concentrated in the FHA loan program and that relatively few minorities receive conventional conforming loans.

loans will approve a higher percentage of the applicant pool than lenders offering lower cost loans. Since high risk applicants have more to lose than low risk applicants if their loan application is denied, they will apply for the loan where they have the best chance of being approved. High risk applicants will apply for high risk loans and low risk applicants will apply for low risk loans.

Since lenders can't distinguish between the applicants with intermediate credit scores, they rely on any factors that will provide additional information about the probability of the applicant being a high risk applicant or a low risk applicant. If the lenders believe that relatively more high risk applicants are concentrated among minority applicants, the race of the applicant may be used as an additional signal of the applicant's creditworthiness. Lenders offering high cost loans will approve a higher percentage of loans to minorities because they are receiving a higher return for making the riskier loans. Lenders offering low cost loans will use race as a signal and deny a high proportion of minority applicants.

The result of minorities having high denial rates in the low cost loan category is that many low risk minority applicants are denied loans while many high risk minority applicants are approved loans. The rejection of loans to low risk minority applicants results in a high denial rate for minority applicants. The concentration of lending to high risk minority borrowers results in a high default rate. This model is consistent with the empirical evidence, which indicates that minority applicants have higher denial rates than white applicants and that minority borrowers have higher default rates than white borrowers. Calem and Stutzer (1995) argue that higher default rates among minority borrowers may be a direct result of this form of discrimination.

Calomiris, Kahn, and Longhofer (1994) also develop a model where higher default rates for minority borrowers is consistent with discrimination against minorities in the mortgage market. In their model, discrimination occurs because information gathering is costly and lenders have a cultural affinity to white applicants. This cultural affinity to white applicants makes it easier to assess the credit risk of white applicants. The model implicitly assumes that the determinants of credit risk may be different for white borrowers than for minority borrowers. In order to accurately assess the credit risk of minority borrowers, lenders must obtain additional costly information.

Assume lenders can invest in signal technology s or t. Signal technology s is more efficient in estimating the credit quality of white applicants, W. Signal technology t is more efficient in estimating the credit quality of minority applicants, B. Both signal technologies are costly for lenders to obtain.

Calomiris, *et. al.* (1994) argue that lenders are more likely to obtain signal technology s which makes them more efficient in underwriting loans to the majority of applicants, W. When underwriting loans to minority applicants, B, lenders must decide whether to use signal technology s or to invest additional funds to obtain signal technology t. If the expected additional profits from underwriting minority loans with signal technology t are less than the cost of acquiring signal technology t, lenders will use signal technology s to underwrite all loans.

The result of using signal technology s to underwrite all loans is that more underwriting errors occur in evaluating minority applicants than white applicants. Since lenders know they will make more underwriting errors in evaluating minority applicants they may be more conservative in investing additional resources in processing marginal loan applications. They rely more on objective and inexpensive information to make loan decisions on minority applicants. Therefore, it is likely that a minority applicant and a white applicant with identical objective criteria will experience different loan outcomes. If the lender is more willing to invest in additional information to support the white applicant's loan application, the white applicant will be more likely to be approved for the loan than the minority applicant. Furthermore, if the lender believes that credit risk is correlated with race, the lender may use the applicant's race as an inexpensive proxy for creditworthiness. Calomiris *et. al.* (1994) show two consequences of using signal technology s to underwrite all loans instead of investing additional financial resources to obtain signal technology t: (1) minority applicants will have higher denial rates than white applicants, because lenders are more willing to consider compensating factors when underwriting white applicants and (2) the average probability of default is greater for minority borrowers because more underwriting errors for minority applicants will result in rejecting some of the creditworthy minorities and approving some of the uncreditworthy minority applicants. This results in lowering the quality of approved minority borrowers, and increasing default rates for minorities.

Ferguson and Peters (1995, 1997) also develop models where higher minority default rates may be consistent with discrimination.[11] Ferguson and Peters (1995) focus on the expected relationships of relative denial and default rates for minority and nonminority borrowers if the distribution of credit quality is higher for nonminorities than for minorities and lenders have perfect information. Ferguson and Peters (1997) extend the model to show the impact of underwriting errors due to noise in the estimation of credit quality and the effect of cultural affinity by lenders as proposed by Calomiris *et. al.* (1994).

Ferguson and Peters (1995) develop a model of lending where the screening process results in a "credit score." The probability the applicant will repay the loan is assumed to be a monotonically increasing function of this score, θ. For simplicity, they assume θ is the probability of repayment. A uniform nondiscriminatory lending policy will result in approving all loans where the creditworthiness of the borrower is above θ^* and denying loans if the score is below θ^*. All borrowers are charged the same interest rate. Assuming each loan is for \$1, the bank receives $(1 + r)$ if the borrower repays the loan. The expected profit on the loan is $\theta r - (1 - \theta)$. The lender selects θ^* such that the profit on a marginal loan is zero.

Assume white applicants have a probability density function , $g(\theta)$, and a cumulative distribution function, $G(\theta)$, and black applicants have a probability density function, $h(\theta)$, and a cumulative distribution function, $H(\theta)$. Further, assume the

[11] The framework for the Ferguson and Peters (1995) model is majority and minority populations, not specifically white and black borrowers. They do, however, indicate their analysis can be clarified by thinking in terms of racial groups.

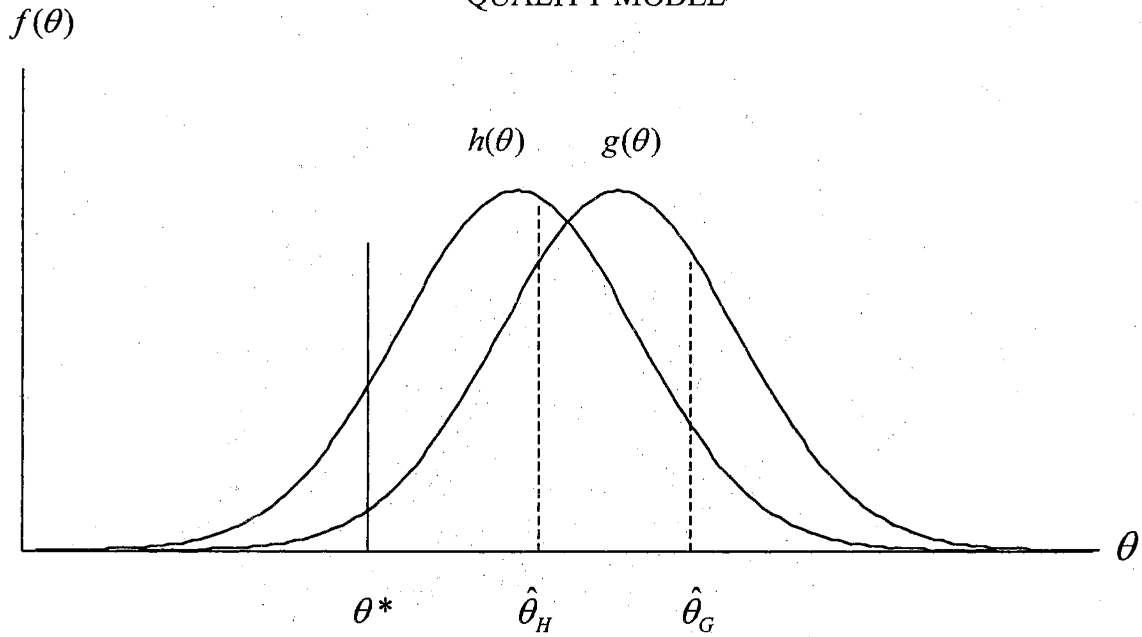
cumulative distribution function, $G(\theta)$, first-order stochastically dominates the cumulative distribution function, $H(\theta)$. Figure 1 shows the probability density functions $h(\theta)$ and $g(\theta)$. Given θ^*, the marginal credit score for all applicants, the average creditworthiness of black borrowers is $\hat{\theta}_H$ and the average creditworthiness of white borrowers is $\hat{\theta}_G$. The marginal white borrowers and the marginal black borrowers have identical creditworthiness, but the average creditworthiness of white borrowers is higher than the average creditworthiness of black borrowers. From Figure 1, it is obvious that black borrowers will have higher denial rates than white borrowers under a uniform credit policy where all applicants with creditworthiness above θ^* are approved and all applicants with creditworthiness below θ^* are denied.[12] Also, because $G(\theta)$ first-order stochastically dominates $H(\theta)$ and $\hat{\theta}_H < \hat{\theta}_G$, black borrowers will have higher default rates than white borrowers.[13] Within this framework, Ferguson and Peters (1995) argue only two combinations of loan denial and default rate experiences can be viewed as discrimination. First, if white applicants exhibit lower denial rates and higher default rates than black applicants, there is evidence of discrimination against black applicants. Second, if white applicants have higher denial rates and lower default rates than black applicants, there is evidence of discrimination against white applicants. In a

[12] The area under the curve to the left of θ^* represents denied applications. The area under the probability density function, $h(\theta)$, to the left of θ^* is greater than the area under the probability density function, $g(\theta)$ to the left of θ^* ($\int_0^{\theta^*} h(\theta)d\theta > \int_0^{\theta^*} g(\theta)d\theta$).

[13] The average default rate for black borrowers is $1 - \hat{\theta}_h$ and the average default rate for white borrowers is $1 - \hat{\theta}_g$.

FIGURE 1

FERGUSON AND PETERS (1995) HETEROGENEOUS CREDIT QUALITY MODEL



nondiscriminatory lending environment, Ferguson and Peters (1995) argue black applicants will have higher loan denial rates and higher loan default rates than white applicants. Observing that blacks have higher denial rates and higher default rates is not evidence of discrimination, reverse discrimination, or no discrimination. Ferguson and Peters (1995) state that no inferences can be made about discrimination under these conditions.

Shaffer (1996) expands the theoretical analysis by Ferguson and Peters (1995) to allow unequal recovery rates on defaulted loans between white borrowers and black borrowers. Shaffer (1996) shows that, if recovery rates are significantly lower on defaulted loans made to black borrowers than on defaulted loans made to white borrowers, a profit-maximizing lending program may result in black borrowers having lower average default rates than white borrowers.

The author's focus on recovery rates is a natural theoretical extension of Ferguson and Peters (1995), but its applicability to the residential loan market is questionable. There is no strong evidence that recovery rates are lower for black borrowers than for white borrowers. All Federal Housing Administration (FHA) loans require government mortgage insurance and it is standard practice to require private mortgage insurance on conventional loans with loan to value ratios exceeding 80 percent. Tootell (1996) finds black borrowers are more likely to be required to get private mortgage insurance than identically qualified white borrowers. This evidence indicates white borrowers are more likely to be allowed to carry high loan-to-value ratios without private mortgage insurance than black borrowers. Therefore, it is possible that recovery rates may actually be slightly lower on defaulted loans made to white borrowers than on defaulted loans made to black borrowers.

Ferguson and Peters (1997) investigate the impact of underwriting errors on marginal loan denial rates when the distribution of credit quality or the underwriting errors differ across racial groups. The authors assume all errors occur at the margin where $\hat{\theta}$ is the minimum credit quality for loan approval and marginal borrowers have credit quality ranging from $\hat{\theta} - \varepsilon$ to $\hat{\theta} + \varepsilon$.[14] All applicants with true credit quality below $\hat{\theta} - \varepsilon$ are correctly denied credit while all applicants with true credit quality above $\hat{\theta} + \varepsilon$ are correctly granted loans. Therefore, all underwriting errors occur in dealing with marginal applicants. Ferguson and Peters (1997) also assume that underwriting errors are symmetric.[15]

Ferguson and Peters (1997) show that random underwriting mistakes will not impact white and minority borrowers to the same degree unless (1) credit quality is homogeneous, and (2) underwriting errors are not correlated with race. If either of these two conditions fails to hold, random underwriting mistakes may impact minorities disproportionately.

[14] Ferguson and Peters (1997) use different notation than Ferguson and Peters (1995). Ferguson and Peters (1995) use θ^* as the marginal cutoff point and $\hat{\theta}$ as the average credit quality for approved borrowers. Ferguson and Peters (1997) use $\hat{\theta}$ as the marginal cutoff point and $\bar{\theta}$ as the average credit quality of all applicants, not just approved borrowers.

[15] A symmetric distribution of errors implies that a lender is equally likely to understate or overstate an applicant's true creditworthiness.

This analysis assumes that the majority of white applicants and the majority of nonwhite applicants meet the minimum requirements for loan approval (i.e. $\bar{\theta} > \hat{\theta}$). The result is that the marginal credit score is on the upward sloping side of the distribution for both white and minority borrowers (Figure 2). Therefore, if underwriting errors occur in a random manner (and are symmetrical as Ferguson and Peters assume), the errors will result in a higher proportion of creditworthy applicants being denied mortgages than the proportion of uncreditworthy applicants that are approved.[16] If the distribution of minority borrowers is lower than the distribution of white borrowers (i.e., white borrowers have a higher distribution of credit quality), then minorities will be more adversely affected by the underwriting errors than white borrowers.[17]

Ferguson and Peters (1997) also argue that lenders may have a cultural affinity to white applicants. This cultural affinity refers to a higher level of efficiency in assessing the credit quality of white borrowers due to having more experience evaluating white borrowers in the past. Even if credit quality is homogeneous, if lenders make more underwriting errors on minority applicants than on white applicants the errors will have a

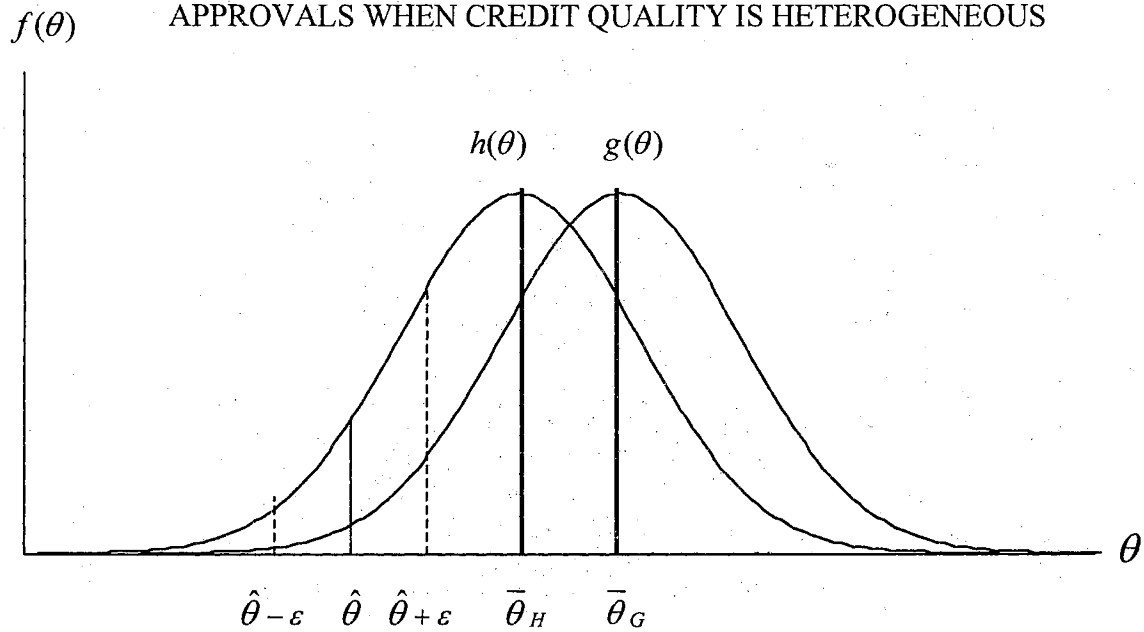
---

[16] The number of creditworthy minorities denied loans will be greater than the number of uncreditworthy minorities granted loans. Assuming p is the proportion of underwriting errors, the number of uncreditworthy minorities granted loans is $\int_{\hat{\theta}-\varepsilon}^{\hat{\theta}} ph(\theta)d\theta$. This is less than the number of creditworthy minorities denied loans, $\int_{\hat{\theta}}^{\hat{\theta}+\varepsilon} ph(\theta)d\theta$.

[17] This is true because $\int_{\hat{\theta}}^{\hat{\theta}+\varepsilon} ph(\theta)d\theta - \int_{\hat{\theta}-\varepsilon}^{\hat{\theta}} ph(\theta)d\theta > \int_{\hat{\theta}}^{\hat{\theta}+\varepsilon} pg(\theta)d\theta - \int_{\hat{\theta}-\varepsilon}^{\hat{\theta}} pg(\theta)d\theta$.

FIGURE 2

THE IMPACT OF RANDOM UNDERWRITING ERRORS ON LOAN APPROVALS WHEN CREDIT QUALITY IS HETEROGENEOUS



disparate impact on minority applicants (Figure 3). That is, minority applicants will have higher denial rates than white applicants due to the underwriting errors.[18]

Longhofer (1996) extends the cultural affinity hypothesis of Calomiris *et. al.* (1994). The Longhofer (1996) model assumes lenders accept applications from two groups, W and M. The lenders are assumed to have a cultural affinity with group W. Lenders receive a signal, s_1, of the borrower's true creditworthiness, θ. The signal can be observed by outsiders. Due to their cultural affinity, lenders receive a second, private signal, s_2, for group W. The second signal is a random signal, and is equally likely to contain information damaging to the loan applicant or information beneficial to the loan applicant. Using this model, Longhofer (1996) shows that the relative denial rates for group W and group M borrowers depends on whether the cutoff point, q*, is higher or lower than the average creditworthiness of the two groups.[19] If q* is above the average creditworthiness of the applicant pool, $\bar{\theta}$, then group M applicants will have higher denial rates than group W applicants.[20] Conversely, if q* is below the average, creditworthiness of the applicant pool, $\bar{\theta}$, group W applicants will have higher denial

---

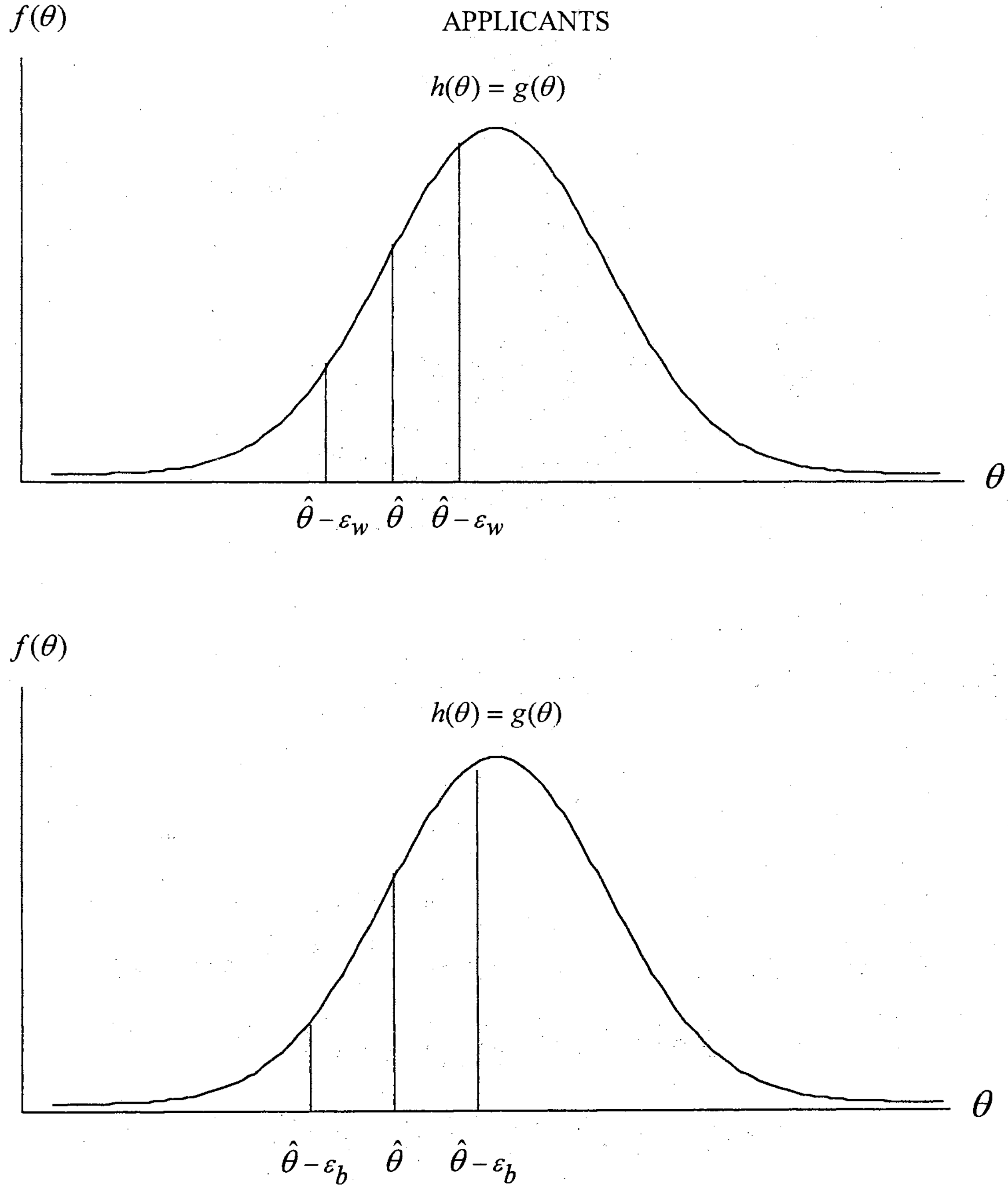
[18] This is true because the marginal cutoff point is on the upward sloping side of the distribution. Assume p_G is the proportion of underwriting errors on white applicants and p_H is the proportion of underwriting errors on black applicants, ($p_H > p_G$). Black applicants will have higher denial rates than white applicants since $\int_{\hat{\theta}}^{\hat{\theta}+\varepsilon} f(\theta)d\theta > \int_{\hat{\theta}-\varepsilon}^{\hat{\theta}} f(\theta)d\theta$.

[19] Here, Longhofer (1996) assumes that group W and group M applicants have the same distribution of credit quality.

[20] This would result in the majority of loan applications being denied, which is contrary to empirical evidence.

FIGURE 3

THE IMPACT OF UNDERWRITING ERRORS ON LOAN APPROVALS WHEN CREDIT QUALITY IS HOMOGENEOUS BUT LENDERS MAKE MORE UNDERWRITING ERRORS ON MINORITY APPLICANTS



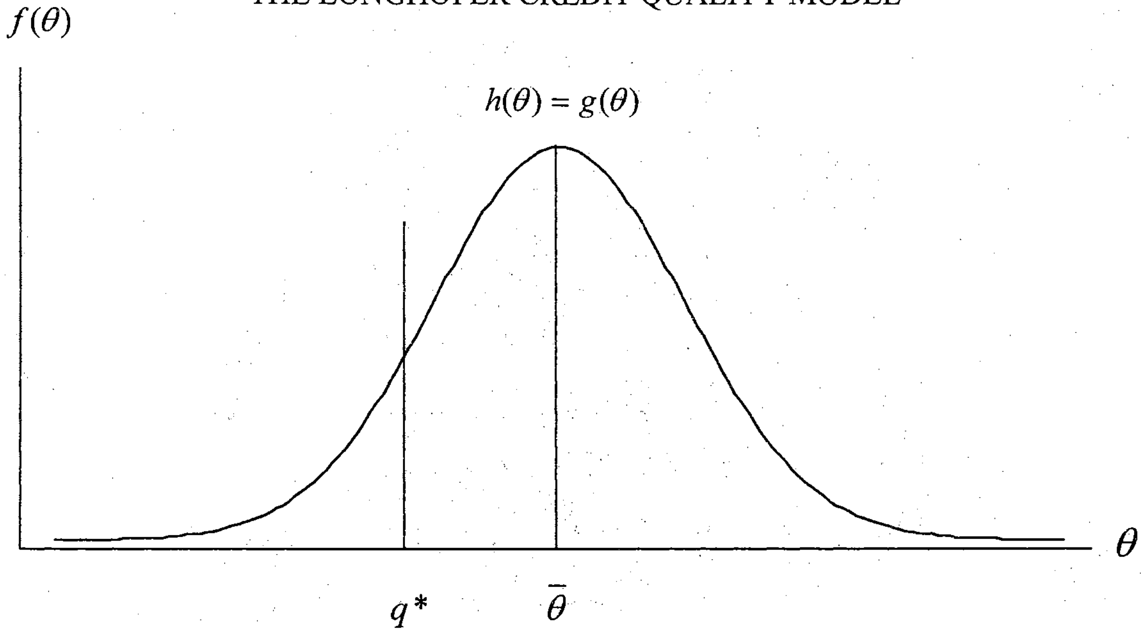
rates than group M applicants. This result is obtained by evaluating the effect of the second signal, s_2, on the loan outcome. If $q^* < \bar{\theta}$, the critical cutoff point is on the upward sloping side of the distribution (Figure 4). Since the area under the curve to the right of q* is greater than the area under the curve to the left of q*, the second signal will result in more negative overrides of loan decisions than positive overrides. That is, group W borrowers will have more loans denied where $s_1 > q^*$ than they will have loans approved where $s_1 < q^*$. The second signal results in group W borrowers having higher denial rates than group M borrowers.

The result shown above applies if the distribution of credit quality is the same for white and minority applicants and if lenders act on additional negative information given in signal, s_2, to the same degree that they act on additional positive information given in s_2. If the distribution of credit quality is sufficiently higher for white borrowers than for minority borrowers, minorities will have higher denial rates than white borrowers. White borrowers will still have more negative overrides than positive overrides, but the lower credit quality of minority applicants will result in minorities having higher denial rates.

The additional assumption of a secondary market with standard loan underwriting guidelines results in a model of cultural affinity that is both realistic and consistent with empirical findings. Longhofer (1996) refers to the Fannie Mae guidelines that outline minimum standards of credit quality, but allow lenders to use compensating factors to approve loans that fall short of the minimum standards. Think of the minimum Fannie Mae underwriting standards as signal, s_1. The lender is more likely to collect additional

FIGURE 4

THE LONGHOFER CREDIT QUALITY MODEL



information on white applicants than minority applicants, but will not collect additional information if the applicant already meets the minimum criteria. Even if the lender does collect negative information on an applicant, they may not use the information in underwriting the loan since the loan will be sold in the secondary market. Since the loan meets the objective minimum underwriting standards of Fannie Mae, the lender has no incentive to use the negative information.[21] However, the lender will use any positive information as a compensating factor in the loan decision. Therefore, the additional information collected on white applicants due to the lender's cultural affinity for white applicants can only lead to positive overrides of the loan decision. In this framework, minorities will have higher denial rates than white applicants even if the distribution of credit quality for the two groups is the same. If the distribution of credit quality for white applicants is higher than the distribution of credit quality for minority applicants, the higher denial rate for minority applicants is compounded even more.

B. *Two Basic Approaches to Empirical Research on Credit Discrimination*

Schill and Wachter (1993) divide credit discrimination research into two categories: aggregate investigations and accept/reject research. Aggregate research investigates the flow of mortgage activity into neighborhoods. These investigations use census data and some Home Mortgage Disclosure Act (HMDA) data to determine if there

---

[21] Longhofer (1996) points out that the loan officer might not even make negative information about the applicant known to the loan committee if the loan already conforms to secondary market underwriting guidelines. To do so might jeopardize the loan and cost the loan officer a commission.

is evidence of redlining in certain areas. Aggregate research on discrimination requires independent variables that measure (1) the demand for mortgage credit, (2) the risk of loss to the lender, and (3) the redlining variables (i.e., the proportion of minorities in the census tract).

Accept/reject investigations use HMDA data, or more detailed data in some cases, to determine if there is evidence of lender discrimination against protected groups on a case by case basis. They employ independent variables to measure (1) the risk of loss to the lender and (2) the redlining variables. Credit decision (accept/reject) investigations do not require measurement of the demand for mortgage credit.

1. Investigations of Aggregate Discrimination

Prior to 1990, most published research on discrimination in the home mortgage lending market were aggregate studies of redlining. Redlining occurs when lenders refuse to make mortgage loans in certain geographic areas. Most redlining research investigates the allocation of mortgage money across census tracts. The early investigations of redlining focus on the supply of mortgage money across neighborhoods without considering the demand for mortgage loans. Canner and Smith (1991, 1992), Canner, Passmore and Smith (1994) and Canner and Passmore (1995) summarize lending patterns using HMDA data. They report that inner-city areas and neighborhoods with a high proportion of blacks receive fewer and smaller loans than suburban areas and neighborhoods with a high proportion of white residents.

Research that considers the demand for mortgage loans across census tracts indicates that much of the disparity in lending is explained by differences in the demand for mortgage credit.[22] Leahy (1985) is one example of an investigation that considers the demand for residential mortgage loans. Leahy (1985) investigates the number and amount of conventional mortgage loans made by the 12 largest lenders in Summit County, Ohio in 1980 to determine if there is evidence of redlining. Leahy (1985) defines the amount of credit flowing into a neighborhood as a function of (1) the demand for loans and (2) selection among applicants for credit. Leahy (1985) uses six variables to measure the demand for conventional mortgage loans in a particular census tract: (1) total population in 1980, (2) the number of owner-occupied units in 1980, (3) the median age in 1980, (4) the percentage of the tract engaged in professional occupations in 1980, (5) the percentage change in the population from 1970 to 1980, and (6) the percentage change in owner-occupied units from 1970 to 1980. Five variables are used to measure the selection among applicants for credit: (1) average number of persons per household in 1980, (2) median home value in 1980, (3) median family income in 1980, (4) percentage of female-headed households in 1980, and (5) percentage of housing units built before 1940.

Leahy (1985) matches black census tracts with white census tracts that are "statistically equivalent" using the eleven empirical variables described earlier in a factor analysis. Leahy (1985) compares the number and amount of conventional mortgage

[22] Leahy (1985), Holmes and Horvitz (1994), and Holmes and James (1996) all find that differences in demand at least partially explain differences in lending patterns to low income and minority neighborhoods.

loans received by each black census tract to its "statistically equivalent" white census tract using an analysis of variance test. Leahy (1985) finds the black census tracts received only 37 percent of the amount of money loaned to the matching white census tracts. Leahy (1985) concludes that prejudice influenced the allocation of mortgage money in Akron, Ohio in 1980. However, Leahy (1985) failed to include any measures to account for home sales within neighborhoods and the volume of home sales is believed to have a significant impact on the demand for mortgage money. Avery (1989) stresses the importance of considering housing transfers in modeling the demand for residential mortgage loans. Even this might not be a good proxy for the demand for mortgage loans because housing transfers would be endogenous to the decision by lenders to redline an area. Housing transfers only represent the demand for loans that has been satisfied. They do not reflect properties that could not be sold because mortgage money was not available to potential buyers.

Holmes and Horvitz (1994) investigate the Houston, Texas mortgage market to determine if the racial composition of census tracts affect mortgage lending activity. They use HMDA data for the years 1988 to 1991 and demographic data from the 1990 census. Holmes and Horvitz (1994) include five general categories in their model: (1) property and loan characteristics, (2) economic condition of the residents in the census tracts, (3) the risk of loans made in each census tract, (4) differences in the mobility of residents, and (5) racial characteristics of the census tracts.

Variables used to proxy property and loan characteristics include (1) median home value, (2) the percentage of vacant homes in the census tract, and (3) the ratio of the average loan to the average home value. The empirical variables used to proxy the economic condition of the residents in the census tract include (1) median household income, (2) percentage of head of households between the ages of 25 and 34, (3) the percentage of adult residents with at least one year of education beyond high school, and (4) the percentage of owner-occupants with mortgage payment to gross income ratios exceeding 30 percent. The risk of loans in the census tract is proxied by (1) the natural log of the ratio of the number of insured loans that defaulted during the period to the number of government-insured loans granted, and (2) the change in median home values from 1980 to 1990. Differences in the mobility of residences are measured by (1) the percentage change in the number of owner occupants from 1980 to 1990, (2) the percentage of residents who have moved into the census tract since 1985, and (3) the percentage of single family homes that are rental units. The dependent variable in the regression is the number of loans made in the census tract expressed as a percentage of the number of owner-occupied homes in the census tract.

Holmes and Horvitz (1994) find census tracts with a high proportion of minorities receive significantly fewer mortgages than census tracts with a high proportion of white residents. However, they conclude differences in mortgage lending across census tracts are justified based on differences in demand for mortgages and in the risk of lending across census tracts. They conclude there is no evidence of redlining in Houston over the period 1988 to 1991. Holmes and Horvitz (1994) point out their research only considers the existence or nonexistence of discrimination across census tracts. It is possible that a disproportionate number of loans in census tracts with large minority populations are made to white borrowers.

Holmes and James (1996) also investigate the Houston market. Holmes and James (1996) investigate the effect of the racial composition of census tracts on the median home values in census tracts. They hypothesize that redlining minority neighborhoods would result in significantly lower median home values in the census tracts being redlined than could be explained by socioeconomic factors. Holmes and James (1996) use a generalized least squares regression model where the dependent variable is the median home value in 1990. The independent variables include (1) the percentage of total population that is black, (2) the percentage of total population that is Hispanic, (3) the change in the percentage of minorities in the census tract between 1980 and 1990, (4) the median household income in 1990, (5) the percentage of the population between 25 and 34 years old, (6) the percentage of adult residents with some education beyond high school, (7) the percentage of single-family homes that are renter occupied, (8) the median age of the housing stock in the census tract, (9) the percentage of single-family homes that are vacant, (10) the loan-to-value ratio, (11) the percentage of owner-occupants with a mortgage payment to gross monthly income ratio greater than 30 percent, (12) the percentage of current owner-occupants who moved in between 1985 and 1990, (13) the percentage change in the number of owner-occupants between 1980 and 1990, (14) the natural log of the default rate in insured mortgages, and (15) the percentage change in median home value.

The authors test for the combined effect of the three race-related variables. They find the race-related variables have a statistically significant impact, but dropping the race-related variables from the regression equation only reduces the R^2 slightly, from .8010 to .7955. The authors conclude race-related variables do not significantly impact home values and lenders in Houston do not redline minority neighborhoods.

Perle, Lynch and Horner (1993) argue that aggregate investigations of mortgage flows cannot identify redlining. They cite problems with four aspects of previous research that limits the usefulness of this type of research: (1) structural modeling, (2) data availability, (3) aggregation level, and (4) model specification. The problem with the structural modeling of previous redlining research stems from the necessity to model both the demand for and supply of mortgages in a neighborhood. A reduced form equation is generated where the dependent variable is a measure of the number or amount of mortgage loans granted in a census tract. The racial composition of the census tract is included in some form in the regression equation as an independent variable. Most redlining research interprets a significant coefficient on the race variable as evidence of redlining. These investigations assume any race effect is related to the supply side of the equation. Perle, Lynch and Horner (1993) state that an implicit assumption of redlining research is that there is no difference in the demand for mortgage loans across racial groups, holding other demand-related variables constant. For this condition to hold, it is a necessary condition that all economic and social variables correlated with race that affect mortgage demand be included in the reduced form equation.

This leads directly to the problem of data availability. Most aggregate research uses HMDA and census data. The data used in these investigations omit several variables that are correlated with race and important in the mortgage lending decision, such as credit history information and total obligations to income ratios.

The aggregation level is also a problem with aggregate redlining investigations. If some lenders in a given market discriminate and others do not, Galster (1993) argues that minority applicants will eventually learn which lenders do not redline and will apply to those lenders for mortgage credit. The result is that redlining will not be found even though some lenders in the area discriminate.

Perle, Lynch and Horner (1993) list three problems with model specification in aggregate investigations of redlining: (1) which functional form should be used, (2) which variables should be included in the model, and (3) the selection of a dependent variable. Perle, Lynch and Horner (1993) investigate the impact of different functional forms, independent variables, and dependent variables on the race coefficient with 1982 mortgage data from the Detroit area and 1980 census data. They test a model where the quantity of mortgage loans is a function of (1) race, (2) socioeconomic characteristics, (3) neighborhood and housing characteristics, and (4) mobility. The percentage of black households is used as the race variable. The percentage of family households, the percentage of female-headed households, and the median household income are used to proxy socioeconomic characteristics. The neighborhood and housing characteristics are proxied by six variables: (1) percentage of the population aged 25 to 34, (2) percentage of houses built before 1939, (3) percentage of houses built between 1960 and 1980, (4) the median value of owner-occupied housing units, (5) percentage of owner-occupied housing units, and (6) percentage of vacant housing units. Mobility is measured by the percentage of persons living in the same house since 1975. The dependent variable used in the analysis is the ratio of the number of conventional and governmental loans to the total number of housing units.

Perle, Lynch and Horner (1993) test the model using both stepwise and direct-entry regression models. They also test the models using linear and log functions and test the eleven variable model against a simpler four-variable model. The four variables in the simple model are (1) median household income, (2) percentage of vacant housing units, (3) percentage of owner-occupied housing units, and (4) percentage of black households. Their results indicate that the log function has slightly more explanatory power than the linear function and the eleven variable model has significantly more explanatory power than the simple four variable model. Stepwise regression did not provide any benefits over the direct-entry regressions.

Perle, Lynch and Horner (1993) find a significant race coefficient with the simple model, regardless of the functional form of the model. They find the race variable to be insignificant in the full model using a log or linear function. The race variable is significant using a multiplicative model, but with a positive sign. The authors conclude that existing aggregate investigations are unsuited for research on neighborhood redlining and encourage the accept/reject methodology for investigations into discrimination in residential mortgage lending.

Phillips-Patrick and Rossi (1996) propose a simultaneous equations approach to the investigation of redlining. They use 1990 census and 1992 HMDA data for the Washington, DC, metropolitan statistical area (MSA) to compare the results of single equation models with the results from a simultaneous equations approach. They estimate two single equation models and one simultaneous equations model in their investigation. The first model includes median income, median home price, and the percentage of residents in the census tract that are black. The coefficient on the race variable in this model is significant at the .01 level. Phillips-Patrick and Rossi (1996) point out that, without further investigation, the conclusion would be that redlining exists in Washington, DC. However, the first model omits several census tract variables that might explain lending discrepancies across census tracts.

The second model has ten independent variables: (1) percentage of black residents in the census tract, (2) median income, (3) median loan-to-value ratio, (4) census tract vacancy rate, (5) median age of residents, (6) unemployment rate, (7) percentage of properties that are rental units, (8) percentage of boarded-up housing, (9) percentage of in-migration in the census tract from 1989 to 1990, and (10) loan amount. The coefficient on the race variable in this model is not statistically significant. Conclusions from this model would be that redlining does not exist in Washington, DC.

The third model estimates a simultaneous equations model. The demand equation includes all the variables from the second model except median loan-to-value ratio and loan amount. The demand equation also includes median home price. The supply side equation includes all of the variables from the second model except the percentage of rental units, the percentage of in-migration from 1989 to 1990, and the median age. The supply side equation also includes the median home price. The demand for loans is measured as the ratio of mortgage applications to total salable units and supply of loans is measured as the ratio of total originations to total salable units. The authors find that loan originations drop in neighborhoods as the percentage of black residents increases. They conclude this could be the result of redlining or it could be the result of an omitted variable problem. Phillips-Patrick and Rossi (1996) point to the different findings of the three models as evidence that redlining investigations using aggregate loan data are inconclusive and can be misleading.

2. *Investigations of Discrimination in Individual Loan Decisions*

Accept/reject investigations have several advantages over aggregate research. First, they don't require estimates of the demand for mortgage loans. Second, they are a more direct test of discrimination in the mortgage markets. Third, the lender(s) involved in discriminatory lending practices may be identified using accept/reject investigations.

Avery, Beeson and Sniderman (1992) use 1990 HMDA data to investigate variations in minority loan originations across lenders. The authors attempt to determine if differences in minority loan originations are primarily due to differences in minority application rates or differences in minority approval rates. The final sample consists of 1,984,688 applications from 8,745 institutions operating in 40,008 census tracts in all 340 metropolitan statistical areas (MSA's). The authors estimate a linear regression equation where the dependent variable is the ratio of minority approvals to total approvals for the lending institution, and the independent variables are the ratio of minority applications to total applications and the ratio of minority approval rates to total approval rates.

Avery, Beeson, and Sniderman (1992) find the variance across lenders in minority and low-income loan originations is accounted for primarily by the variance in application rates as opposed to differences in the disposition of the applications. For the U.S. as a whole, 87 to 91 percent of lender-specific differences in credits to minorities were accounted for by differences in minority application rates. These results held for all types of institutions, for different lender sizes, for different lender market shares, and for various definitions of the relevant market.

Avery, Beeson, and Sniderman (1992) also investigate the differences in minority and nonminority denial rates. They employ a linear probability model with the following independent variables: (1) income, (2) a dummy variable for the gender of the applicant, (3) a dummy variable to indicate if the loan application is for a government-insured loan, (4) loan amount, (5) a dummy variable to indicate the MSA, (6) a dummy variable to indicate the census tract, and (7) a dummy variable to indicate the lender. Avery, Beeson, and Sniderman (1992) find denial rates for minority loan applications cannot be explained by the applicants' economic characteristics. Minorities have a 25.2 percent denial rate while nonminorities have a 13.1 percent denial rate. The authors conclude the unexplained differences in denial rates may result from lender bias or from differences in the unobserved characteristics of the loan application. The analysis excludes variables such as credit and employment histories, loan-to-value ratios and total obligations-to-income ratios which are believed to influence the loan decision.

A second analysis by Avery, Beeson and Sniderman (1993b) uses 1990 HMDA data to investigate racial differences in lending. The investigation addresses the possibility that differences in minority loan approval rates can be explained by applicant characteristics other than race or property location. The linear probability model employed estimates the probability that a random loan application will be denied based on applicant characteristics, race, MSA, census tract, and lender. The investigation includes loans for the purpose of home purchase, refinance, and home improvement. The investigation reveals denial rates for minority applicants are consistently higher than those for white applicants with otherwise identical attributes as reported in the HMDA data. The conclusions are consistent across geographic markets and loan products. The evidence suggests racial differences in denial rates are widespread and cannot be attributed to particular areas or certain types of lenders.

Schill and Wachter (1993) investigate discrimination in mortgage lending using 1992 HMDA and 1990 census data for Boston and Philadelphia. The authors specify a logit regression model where the probability that a loan is approved is a function of individual loan characteristics and census tract characteristics. Applications for home purchase, home improvement, and refinancing are tested separately.

The dependent variable in the logit regression equation equals one if the loan was approved and zero if it was rejected. Schill and Wachter (1993) specify two models. The first model has ten independent variables derived from the HMDA data. These variables measure individual loan characteristics. One of the independent variables is the percent of households in the census tract headed by a person who is black. The second model includes the ten independent variables from the first model and adds seven more independent variables from census data to measure census tract characteristics.

Schill and Wachter (1993) find the racial composition of the census tract is statistically significant in the first model and not statistically significant in the second model. The results hold in Boston and Philadelphia for all loan products except applications for mortgage refinancing in Boston. In this case, the race coefficient is negative and statistically significant at the .05 level, indicating black applicants were less likely to have applications for mortgage refinancing approved in Boston than identically qualified white applicants. Schill and Wachter (1993) conclude the weight of the evidence is inconsistent with theories of redlining against minority neighborhoods. The authors find neighborhood risk factors, such as percent of owner-occupied homes, percent of vacant housing units, and median household income help explain differences in mortgage flows across neighborhoods.

Schill and Wachter (1993) emphasize the need for a more complete dataset to come to any definitive conclusions about redlining. The authors also emphasize further research is needed to determine the sources of any racial or ethnic geographic disparities in mortgage lending.

A recent investigation by the Federal Reserve Bank of Boston employs the most comprehensive loan application information of any of the recent credit decision investigations (Munnell, Tootell, Browne, and McEneaney (1996)).[23] The sample is

[23]This investigation was conducted and first released in 1992 as Boston Federal Reserve Working Paper No. 92-7.

taken from the 1990 HMDA disclosures of financial institutions in the Boston MSA. The sample of applications exhibits denial rates for white applicants of 10.3 percent and denial rates for black/Hispanic applicants of 28.1 percent. The data set is expanded to include 38 additional variables that might influence the loan decision (e.g., net wealth, loan-to-value ratios and total debt payments to income ratios).[24]

Munnell *et. al.* (1996) assume mortgage lenders maximize profits by minimizing the probability and costs of default associated with each mortgage loan. The probability of a lender denying a mortgage application is hypothesized to be a function of (1) the applicant's ability to meet loan payments, (2) the risks of default, (3) the potential loss associated with default and foreclosure, and (4) the terms of the loan. The model also considers personal characteristics, such as race. The ability of the applicant to meet loan payments is measured by obligation ratios and net wealth. The risk of default is measured by credit histories, public record histories of credit problems, and employment characteristics. The potential loss associated with default is measured by the loan-to-value ratio, a dummy variable for denial of private mortgage insurance, and a dummy variable indicating if the loan will be secured by a two-to-four family dwelling. These variables are included in a logit regression equation with a dummy variable for the race of the applicant. The dependent variable is the probability that the mortgage loan application will be denied. Munnell *et. al.* (1996) run several models using different

[24] The authors indicate they had extensive conversations with lenders, underwriters, and examiners and the 38 additional variables collected include any variables considered to be important by these groups.

empirical variables to find the model with the best "fit." The race coefficient is highly significant in every model specification tested.

Munnell *et. al.* (1996) find minority applicants have greater debt burdens, higher loan-to-value ratios, weaker credit histories, and are less likely to purchase a single-family home than white applicants. These factors account for a large portion, but not all, of the disparity between minority and white denials. Black and Hispanic applicants are rejected 60 percent more often than white applicants when financial, employment, and neighborhood characteristics are held constant. Munnell *et. al.* (1996) also find denied black/Hispanic applicants on average have poorer objective qualifications than denied white applicants. Denied minorities have lower income and wealth, higher obligation and loan-to-value ratios, and worse credit histories than denied whites. The authors conclude Boston area lenders discriminated against minorities, particularly black applicants, in the market for home mortgages in 1990.

The Munnell *et. al.* (1996) investigation has been criticized for errors in data coding, poor model specification, and questionable interpretation of the statistics by the authors. Data coding errors are pointed out by Horne (1994), Liebowitz (1993), and Carr and Megbolugbe (1993). Horne (1994) with FDIC staff, examined the loan files of FDIC-supervised institutions included in Munnell *et. al.* (1996). The FDIC staff reviewed all loans that appeared to be discriminatory. They find no evidence of discrimination, but they do find numerous cases where data had been miscoded by the financial institutions participating in the sample or by the authors. Liebowitz (1993) claims that omitting the observations with inconsistent data or extreme observations results in an insignificant race coefficient. Carr and Megbolugbe (1993) conduct a procedure to identify errors and "clean" the Boston Fed dataset. They find the data errors are not responsible for the significant race effect in Munnell *et. al.* (1996). Their findings with the clean data are very similar to the findings reported in Munnell *et. al.* (1996).

Zandi (1993) criticizes Munnell *et. al.* (1996) for improper model specification. Zandi (1993) claims Munnell *et. al.* (1996) suffers from omitted-variable bias. He finds the impact of race is greatly reduced if four variables are added to the Boston Fed model. The variables Zandi includes are (1) a dummy variable which indicates if the loan meets loan underwriting guidelines, (2) a dummy variable to indicate if some information was unverifiable, (3) a dummy variable indicating if there is a cosigner, and (4) the loan amount.[25]

Tootell (1996) uses data collected by the Boston Federal Reserve in 1992 to investigate the Boston area market for evidence of redlining.[26] The focus of the investigation is to distinguish redlining neighborhoods from the effects of racial discrimination against individual applicants. Munnell *et. al.* (1996) find evidence of discrimination against black borrowers using the Boston Fed dataset. Tootell (1996) investigates the importance of the racial composition of the neighborhood in the lending

---

[25] Browne and Tootell (1995) state the Munnell *et. al.* (1996) research did include the loan amount and a dummy variable indicating the existence of a cosigner and neither variable was significant. Browne and Tootell (1995) argue against including verification and credit standards questions because both "involve an ex post judgment by the respondent." The other variables in Munnell *et. al.* (1996) are based on objective criteria.

[26] Tootell (1996) uses the same dataset used in Munnell *et. al.* (1996). For a more complete description of the dataset and how it was originated, refer to the discussion of Munnell *et. al.* (1996).

decision after controlling for the race of the applicant. Tootell (1996) argues the Boston area market is ideally suited for this type of analysis because over 50 percent of the minority applications received by Boston area lenders in 1992 were for properties located in predominately white census tracts.

Tootell (1996) estimates a logit regression model and a linear probability model where the dependent variable equals one if the loan application is denied and zero otherwise. First, Tootell (1996) estimates a model using the variables used in Munnell *et. al.* (1996), except for the tract characteristic variables. The only tract characteristic variables Tootell (1996) includes are variables to measure the racial characteristics of the census tract. Two variable specifications are used. The first specification uses a dummy variable to indicate if the minority population in the census tract exceeds 30 percent. The second variable specification is the percentage of minority residents in the census tract.

When the race of the applicant is excluded from the models, the variables measuring the racial composition of census tracts are highly significant. When the race of the applicant is included in the models, the race variable is highly significant and the variables measuring the racial composition of the census tract are no longer significant.

Tootell (1996) concludes the racial composition of the census tract does not significantly affect the mortgage lending decision, but the race of the applicant does. The evidence suggests racial discrimination against individual loan applicants, not redlining of minority neighborhoods.

Becker (1993) criticizes investigations of racial discrimination in mortgage lending such as Munnell *et. al.* (1996) for calculating denial rates instead of looking at the profitability of loans to different groups. Becker (1993) and England (1993) argue that, if discrimination exists, minorities should have lower default rates than nonminorities. They contend that failing to observe lower default rates for minority borrowers is evidence against racial discrimination in mortgage lending. Brimelow and Spencer (1993) also use this reasoning to challenge the findings of Munnell *et. al.* (1996). They cite the Boston Fed's finding that the average mortgage default rate for minority neighborhoods in Boston is the same as the rate for white neighborhoods. Brimelow and Spencer (1993) argue equal default rates for minority and white neighborhoods contradicts the Munnell *et. al.* (1996) conclusion that Boston area lenders discriminate against minority applicants.

Munnell *et. al.* (1996), Tootell (1993), Browne and Tootell (1995), Galster (1993), and Ferguson and Peters (1995) argue that racial discrimination in the mortgage market will result in lower default rates for minority borrowers only if certain conditions hold. Tootell (1993) and Browne and Tootell (1995) argue equal minority and white default rates can only be used as evidence of nondiscrimination if the distribution of the quality of accepted minority applicants is identical to the distribution of accepted white applicants in a nondiscriminatory lending framework.[27] Galster (1993) argues

[27]In a nondiscriminatory lending framework where the quality of black and white applicants are identically distributed, default rates would be equal for the two groups. Any discriminatory behavior that involved holding minorities to a higher standard than white borrowers would result in lower default rates for black borrowers. This assumption is critical for the arguments of Becker (1993) and England (1993) to be valid.

inequalities between whites and minorities in occupations, income, indebtedness, and assets will result in minority mortgage holders being distributed more heavily in higher default risk categories. He argues minority borrowers will have higher default rates than white borrowers in a nondiscriminatory lending environment.

LaCour-Little (1996) uses reverse regression on the data from Munnell *et. al.* (1996) to test for discrimination. First, the author estimates a logit regression equation with eleven independent variables used in Munnell *et. al.* (1996). The dependent variable, ACTION, equals one if the loan was denied. There is no race coefficient in this model. Predicted probabilities of loan denial are estimated for each observation. The predicted probabilities are considered the inverse qualifications index, Q-INDEX. Each loan has a Q-INDEX, an estimated probability of loan denial. The Q-INDEX values are then used as the dependent variable in an ordinary least squares regression. The independent variables in this regression are the action taken by the lender, ACTION, and the race of the applicant, RACE. The purpose of this procedure is to determine, on average, whether minority applicants are more or less qualified than white applicants given that the loan was approved. A value, a^*, is calculated by taking the negative of the RACE coefficient and dividing by the ACTION coefficient. This value, a^*, is a measure of the average qualifications of accepted minority applicants relative to accepted white applicants. The value of a^* in the reverse regression procedure is -.193. LaCour-Little (1996) interprets this value to indicate that accepted minority applicants had average qualifications 19 percent lower than accepted white applicants. LaCour-Little (1996) concludes that lenders hold white applicants to higher standards than minority applicants and argues this is evidence of reverse discrimination.

The fact that approved minority applicants have average qualifications 19 percent lower than approved white applicants is not necessarily proof of reverse discrimination as LaCour-Little (1996) contends. The findings simply substantiate what many researchers have argued. The distribution of the quality of accepted minority applicants is lower than the distribution of the quality of white applicants. If the model developed by Ferguson and Peters (1995) holds, and empirical evidence in Munnell *et. al.* (1996) indicates that it probably does, then the finding by LaCour-Little (1996) is to be expected. The findings of LaCour-Little (1996) actually provide further evidence that the distribution of the quality of accepted minority applicants is lower than the distribution of accepted white applicants.

Hunter and Walker (1995) empirically test the cultural affinity hypothesis of Calomiris *et. al.* (1994). According to Hunter and Walker (1995), if the cultural affinity hypothesis is true, lenders will rely more heavily on objective measures of credit quality when evaluating minority loan applicants. Therefore, minority applicants' probability of loan approval should be more sensitive than white applicants' probability of loan approval to changes in these objective measures.

The authors use a subset of the Boston Federal Reserve loan application data used in Munnell *et. al.* (1996). The dataset consists of 1,991 loan applications taken by Boston area lenders in 1990. The dataset includes 1,726 approved loan applications and 265 denied loan applications. There are 1,516 white applicants and 475 black and Hispanic applicants in the dataset. The actual loan approval rate for white applicants in the dataset is 90 percent compared to 76 percent for black and Hispanic applicants.

Hunter and Walker (1995) estimate a logit regression model that includes 26 independent variables. The independent variables are similar to those used in the Munnell *et. al.* (1996) investigation, but also include dummy variables to measure (1) education level, (2) whether or not there is a co-signer on the loan, (3) the thickness of the loan file (the file is considered thick if there are two or more credit reports in the file), and (4) the sex of the applicant.

The authors isolate the affects of each independent variable on the probability of loan approval by holding all the independent variables constant except one and examining the effect of changes in the independent variable on the probability of loan approval. The impact of changes in one independent variable on the probability of loan approval is determined for high quality borrowers and for marginal borrowers by manipulating the values of the remaining independent variables.[28]

Hunter and Walker (1995) find that race is unimportant for high quality applicants. For marginal applicants, credit history problems and high obligations-to-income ratios are more detrimental if the borrower is a minority. Holding other factors constant, marginal white applicants with a total obligations-to-income ratio of .3 have a 94 percent probability of loan approval rate while marginal black applicants with identical objective characteristics have an 88 percent probability of loan approval. If the

[28] For high quality applicants, the independent variables are held constant at values that are consistent with good credit quality. For marginal applicants, the independent variables are held constant at values that reflect significantly greater credit risk.

same marginal white and minority applicants have a total obligations-to-income ratio of .6, the probability of loan approval for the white applicant is 70 percent compared to just 16 percent for the minority applicant.

Based on their findings, Hunter and Walker (1995) conclude that lenders do rely more heavily on objective criteria when evaluating minority loan applicants. This is consistent with the cultural affinity hypothesis of Calomiris *et. al.* (1994).

3. *Current State of the Research on Discrimination in Residential Mortgage Lending*

There is consistent evidence that minorities are two to three times more likely to be denied residential mortgage loans than nonminorities. There is also consistent evidence that predominantly black census tracts receive fewer residential mortgage loans than predominantly white census tracts. The current debate focuses on the reason for these racial disparities. Are these disparities the result of lenders having a taste for discrimination? Are they the result of owner wealth maximizing behavior by lenders? Do lenders practice statistical discrimination by using race as an inexpensive proxy to measure credit risk? Finally, do lenders have a cultural affinity with white applicants that makes them more likely to consider factors other than the objective criteria normally used?

Empirical investigations of aggregate research (i.e., redlining) and accept/reject investigations have come a long way and answered some important questions. For example, well constructed investigations of redlining have consistently failed to find evidence that lenders redline low income neighborhoods. However, there are still many unanswered questions, including whether or not lenders discriminate against minorities in individual credit decisions and which variables should be included as proxies for credit risk in accept/reject investigations.

The variables used in the accept/reject research must accurately reflect the default risk of the borrower. Two types of errors are possible that will lead to incorrect conclusions about whether or not discrimination occurs. First, variables that influence default risk and are highly correlated with race may be omitted from the credit decision model. Several previous investigations have suffered from this omitted variable bias. In these investigations, a finding of discrimination may result from omitting an important variable from the model.

Second, the model could include economic variables highly correlated with race that do not accurately reflect default risk. Including these variables in the credit decision model should lead to a finding of no discrimination. However, if the credit decision model includes variables that do not accurately reflect default risk and the variable results in higher denial rates for minorities, discriminatory lending may be present and not identified.

The methodology employed in this research incorporates the borrower's loan default decision into the credit decision. This analysis is crucial to assure that only variables affecting default risk are included in the credit decision model.

CHAPTER III.
THE EVIDENCE ON LOAN DEFAULT MODELS

The nature of the credit decision process is central to a discussion of racial disparities in mortgage lending. The credit decision focuses on the probability of default by the borrower and the cost to the lender in the event a default occurs. The nature of the credit decision process could produce racial disparities if economic variables correlated with race that do not accurately reflect default risk are included in the credit decision. The determinants of default risk have received extensive analysis. Quercia and Stegman (1992) divide the research on residential mortgage defaults into three categories, which they refer to as first, second, and third generation research.

A. First Generation Research

This line of research views default from the lender's perspective. These investigations attempt to identify loan, borrower, and property characteristics at the time of loan origination that contribute to loan default. These investigations do not tend to be based on well-developed theoretical models. They simply focus on identifying the determinants of residential mortgage defaults.

One of the earliest investigations into the determinants of residential mortgage defaults is von Furstenberg (1969). Using FHA origination data for 1957 through 1965 and default data on the same loans from 1962 through 1966, von Furstenberg (1969) investigates the effect of financing terms on default probabilities. The research involves estimating a multiple regression equation where the dependent variable is the ratio of defaulted loans to total FHA endorsed loans for each year. The independent variables are the equity ratio (measured as one minus the loan-to-value ratio), the duration of the mortgage, and the duration of the mortgage squared:

$$D_t = b_0 + b_1(1 - L/V) + b_2(t) + b_3(t^2).$$

The equation is estimated separately for original maturities of 20, 25, and 30 years and for new and existing homes. A new home is defined as a home not previously owner occupied.

Regression results indicate the equity ratio is negative and significant at the .01 level in each of the six regression calculations. The duration of the mortgage is positive and significant at the .01 level in five of the six calculations, while the duration squared is negative and significant at the .01 level in five of the six calculations. The duration and duration squared are not statistically significant for loans with original maturities of 20 years secured by new homes.

These findings indicate equity reduces default risk and default risk increases as the age of the mortgage increases, but at a decreasing rate. The importance of home equity is highlighted by von Furstenberg (1969). Using a 90 percent loan-to-value ratio as the base case, von Furstenberg (1969) shows that as the loan-to-value ratio increases from 90 percent to 97 percent, default rates quintuple. Loans with loan-to-value ratios of 76-80 percent result in default rates at one-third the level of loans with 90 percent loan-to-value ratios.

These findings lead von Furstenberg (1969) to the conclusion that home equity is the dominant factor affecting default rates. He also finds higher default rates on loans with longer original terms to maturity.

Herzog and Earley (1970) investigate the influence of loan characteristics, borrower-related factors, and economic factors on loan delinquencies and default. The authors use 12,581 FHA and VA loans originated in 1963 to estimate separate regression models for loan delinquencies and loan defaults.

Loan characteristics in the model include the loan-to-value ratio, a dummy variable to indicate the existence of a second mortgage on the property, and the term to maturity. Borrower-related factors in the model are the number of dependents, marital status, payment-to-income ratio, and dummy variables to measure (1) if the borrower is employed in a professional occupation and (2) if the borrower is self-employed. A dummy variable for the region of the country is used to proxy economic conditions.

Herzog and Earley (1970) conclude that the existence of a second mortgage is the most important factor in explaining default. Defaults are positively related to the existence of junior financing. The existence of a second mortgage increases the loan-to-value ratio for the borrower, and reduces the equity the borrower has in the home. Therefore, this relationship is not surprising. Results of the analysis also indicate loans with higher loan-to-value ratios and self-employed borrowers are more likely to default, while borrowers employed in a professional occupation are less likely to default. Contrary to the findings of von Furstenberg (1969), Herzog and Earley (1970) do not find the term to maturity to be a significant factor in explaining loan delinquency or default. The payment-to-income ratio, marital status, and number of dependents are also statistically insignificant in the Herzog and Earley (1970) investigation.

Vandell (1978) develops a general model of default risk where the factors affecting default risk are "borrower related effects, the payment burden over time, the equity accumulation over time, and a transient time term." Vandell (1978) uses borrower income, Y, at time t to proxy borrower-related effects. He hypothesizes that income will be negatively related to default risk and will be more important in lower income ranges. The payment burden is measured as the payment-to-income ratio, Q/Y. Default risk is expected to increase at an increasing rate as the payment -to-income ratio increases. The equity accumulation over time is measured as the contemporaneous equity-to-value ratio, E/Y, where equity includes the downpayment, amortization of the mortgage, and equity obtained through property appreciation. The equity ratio is hypothesized to be negatively related to default risk and changes in equity are expect to influence default risk more at low equity levels than at high equity levels. Vandell (1978) estimates a log-linear model of default risk where:

$$\log D_t = \log C + b_1 \log Y + b_2 \log(Q/Y) + b_3 \log(E/Y) + b_4 \log t.$$

Using the model developed by von Furstenberg (1969) and transforming variables to fit his model using simulations, Vandell (1978) fits the simulated data to his regression equation.[29] Vandell (1978) finds the contemporaneous equity ratio is the dominant factor in explaining loan defaults. The payment-to-income ratio and income variables are not statistically significant.

The dominant finding of first generation research into residential loan defaults is that home equity plays a central role in explaining loan defaults. There is little evidence

[29] The Vandell (1978) model includes some variables not included in the von Furstenberg (1969) model and includes other variables that can change over time. Vandell (1978) makes several assumptions to calculate these variables in the simulations (i.e., constant income and property values, property values are related to income only).

that borrower characteristics such as employment history and payment burdens are important factors.

B. *Second Generation Research*

The second generation research views default from the borrower's perspective. This line of research is based on the economic theory of consumer behavior (Quercia and Stegman (1992)). These investigations model the decision by the borrower to default or continue paying as a decision made by borrowers to maximize their utility over time. In this sense, utility maximization is generally meant to imply wealth maximization.[30]

Jackson and Kaserman (1980) formally develop two competing theories to explain residential mortgage defaults. They refer to these theories as (1) the equity theory of default and (2) the ability to pay theory of default. The purpose of the investigation is to derive testable hypotheses to determine if the empirical evidence supports one theory over the other.

The equity model developed by Jackson and Kaserman (1980) assumes borrowers will default if this course of action results in the best financial outcome for the borrower. The model assumes borrowers will default if the market value of the mortgaged property falls below the outstanding balance on the loan. The probability of default is equal to the probability that the market value of the property falls below the loan balance. Jackson and Kaserman (1980) use comparative statics to determine the expected influence of the

---

[30] Although utility maximization would include other factors, such as pride in homeownership, these factors are not included in investigations of the determinants of default because they are virtually impossible to measure.

loan-to-value ratio, term to maturity, and contract interest rate on the probability of default. According to the equity theory of default developed by Jackson and Kaserman (1980), all three variables should be positively related to the probability of default.

The ability to pay model of default assumes borrowers will continue to make required mortgage payments as long as their current income, net of expenditures that are deemed more important than the mortgage payment, is sufficient to do so. The probability of default is equal to the probability that the income available to make the mortgage payment is less than the required payment. Jackson and Kaserman (1980) conduct comparative statics analysis on the ability to pay model to determine the expected influence of the loan-to-value ratio, term to maturity, and contract interest rate on the probability of default. The only sign that is different for the two models is the term to maturity.

The equity theory of default indicates longer maturities should increase the probability of default. This is reasonable because a longer term to maturity results in slower equity accumulation, making it more likely the market value of the property will fall below the outstanding balance on the loan. The ability to pay theory of default indicates a longer term to maturity should decrease the probability of default. Longer terms to maturity reduce the required payment on the mortgage, making it less likely that the income available to pay the mortgage will fall below the required mortgage payment.

Jackson and Kaserman (1980) use this difference to test which model the data supports. Jackson and Kaserman (1980) estimate regression equations using the loan-to-value ratio, term to maturity, and contract interest rate on the loan with data on 1,736 FHA loans originated in 1969. Results of the empirical tests indicate all three empirical variables are positively related to loan default. Only the loan-to-value ratio is significant at the .05 level. Even though the term to maturity is only marginally significant (at the .10 level), Jackson and Kaserman (1980) conclude the empirical evidence supports the equity theory of default over the ability to pay theory of default.

Campbell and Dietrich (1983) develop a single period model of the default decision where borrowers are assumed to choose the qualitative choice, S, that maximizes their utility. In the Campbell and Dietrich (1983) model, the borrower has four choices: (1) default, (2) become delinquent, (3) prepay the mortgage, and (4) continue payment. Campbell and Dietrich (1983) hypothesize the borrower's current equity position (measured as the contemporaneous loan-to-value ratio) and the current payment to obligations ratio are the principal determinants in the borrower's default decision. This is similar to the development by Jackson and Kaserman (1980) of an equity model of default and an ability to pay model of default. However, Campbell and Dietrich (1983) recognize these two factors may work together to cause default. They argue increasing payment-to- income ratios that make borrowers unable to make their mortgage payments will only lead to default if that is the lowest cost solution. If the borrower has significant equity in the property, default is unlikely regardless of the payment to income ratio. Campbell and Dietrich (1983) argue the probability of default should be positively related to the payment to income ratio and the loan-to-value ratio. They also argue the probability of default should be negatively related to the relative spread between the current mortgage rate and the original mortgage rate. A positive spread makes default less likely because borrowers may be able to allow potential homebuyers to assume their loan, effectively increasing the marketability of their residence. Campbell and Dietrich

(1983) also include (1) the regional unemployment rate, (2) a dummy variable to indicate if the house is new, and (3) the age of the mortgage (and age squared) in their empirical model. The regional unemployment rate is expected to be positively related to the probability of default. Campbell and Dietrich (1983) hypothesize the probability of default will be higher for new homes than older homes because they believe buyers of new homes tend to be younger and more transient and might be more likely to experience personal and financial problems than most borrowers. The age of the mortgage is expected to have a nonlinear relationship with the probability of default. Campbell and Dietrich (1983) utilize a multinomial logit regression equation where the probability of default is the dependent variable.

The data used in the Campbell and Dietrich (1983) research consists of approximately 2.5 million conventional residential mortgage loans originated and insured by the Mortgage Guaranty Insurance Corporation (MGIC) between 1960 and 1980. The MGIC maintains a continuous record of each loan until insurance is canceled or the mortgage defaults. The investigation also uses regional economic data. The economic data used includes indices of mean housing prices and nominal disposable income, the unemployment rate and the current mortgage rate in the region. The data is aggregated each year based on (1) the state where the mortgage was originated (50 states and the District of Columbia), (2) the age of the dwelling (old or new), and (3) the original loan-to-value ratio (four categories).[31] The number of observations in the sample is reduced to 4,899 by eliminating observations with less than 50 individual loans.

[31] For each year, observations are placed in one of 408 groups (51 states times 2 age classifications x 4 loan-to-value ratio categories.

Campbell and Dietrich (1983) run regressions with and without the age of the mortgage and the age of the mortgage squared as independent variables. The current payment to income ratio is positive and significant in the regression that includes the age variables, but is negative and insignificant when the age variables are excluded. The current loan to value ratio, the unemployment rate, and the dummy variable indicating a new residence are positive and significant in both default model specifications. The ratio of the current mortgage rate to the contract rate is negative and significant for both model specifications. Campbell and Dietrich (1983) conclude the current loan-to-value ratio significantly influences the default decision and that unemployment rates had a significant impact on the probability of default in the 1960's and 1970's.

Vandell and Thibodeau (1985) also model the default decision using a model of borrower choice. The model assumes that borrowers make decisions to maximize their wealth in the terminal period. Their model is based on the Campbell and Dietrich (1983) model. In the Vandell and Thibodeau (1985) model, the borrower has five choices: (1) default, (2) become delinquent, (3) prepay through refinancing, (4) prepay through resale, and (5) continue payments. Vandell and Thibodeau (1985) express the payoff function if the borrower defaults as:

$$W_D = (Y - R - Q_r)(1 + r_i) + W(1 + r_i)$$

where,

W_D = expected terminal real after-tax wealth if the borrower defaults,
Y = real annual after-tax household income,
R = required real nondiscretionary expenditures (other than housing),
Q_r = required real rent on new unit,
r_i = expected real return on nonhousing investments,
W = current real nonhousing wealth.

The payoff function if borrowers continue payments is:

$$W_C = (Y - R - Q)(1 + r_0) + (V_T - L_T) + W(1 + r_i)$$

where,

W_C = expected terminal real after-tax wealth if the borrower continues making payments,
Q = required real after-tax payment on mortgage (plus taxes, insurance and other ownership costs),
r_b = expected opportunity cost of borrowing,
r_i = expected real return on nonhousing investment,
$r_0 = r_i$, if $Y - R - Q >= 0$, or $r_0 = r_b$, if $Y - R - Q < 0$),
V_T = expected real market value of current home,
L_T = expected real outstanding loan balance on current mortgage.

Based on the model, borrowers will choose to default if the expected terminal real after-tax wealth from defaulting on the mortgage is greater than the expected terminal real after-tax wealth from continuing to make payments ($W_D > W_C$). Assume real income exceeds real expenditures ($Y - R - Q > 0$). The excess of real income is assumed to be invested in nonhousing investments where it will earn the expected real return on nonhousing investment, r_i. The borrower will default if the equity in the home is expected to be less than the savings from renting plus interest:

$$V_T - L_T < (Q - Q_r)(1 + r_i).$$

One major contribution of the Vandell and Thibodeau (1985) model is the recognition that the rental costs of a new unit must be considered. If a comparable unit can be rented for much less that the required mortgage payment (Q is much greater than Q_r), then default is more likely. The impact of the required rental payment on the probability of default is modeled formally for the first time by Vandell and Thibodeau (1985).

Vandell and Thibodeau (1985) use 28 empirical variables to estimate their model. The empirical variables include three loan-related variables, three financial market variables, 16 borrower-related variables, and six variables proxying housing market and economic conditions.[32] Vandell and Thibodeau use historical data on 348 conventional, owner-occupied loans originated from 1972 to 1983 in the Dallas, Texas area. The model is tested using a logit regression model with 2073 observations, one observation for each year each loan was outstanding. The dependent variable in the analysis is whether or not the borrower defaulted on the loan.

Vandell and Thibodeau (1985) find nonequity factors play an important role in the borrower's decision to default. They find several borrower characteristics influence default. Self-employed borrowers and borrowers working on a commission basis are found to be more likely to default than borrowers with salary income. The length of employment is negatively related to default. Wealth levels are also found to be negatively related to default. These findings are consistent with the theoretical model developed in Vandell and Thibodeau (1985). The authors conclude that nonequity factors play an important role in determining whether or not borrowers exercise their default option. One weakness of the empirical analysis in Vandell and Thibodeau (1985) is the low number of defaulted loans in the sample. Of the 348 loans in the sample, only 37

[32] Borrower-related variables include such things as the marital status of the borrower, whether or not income is commission-based, length of employment, number of dependents, and the sex of the borrower. Loan-related variables include the expected loan balance to market value ratio and the contemporaneous mortgage payment to household income ratio.

defaulted. Of the 2,073 observations created for the empirical tests, only 37 have a dependent variable equal to one.[33]

Giliberto and Houston (1989) develop a theoretical model of the default decision that incorporates relocation costs/benefits. Borrowers are assumed to take actions that will maximize their wealth. Wealth is defined as:

$$W = NW + E + \max[MB_o - MV - RFNC] + h_o,$$

where NW = nonhousing wealth,
E = book value of equity
Mb_o = the current principal balance on the mortgage,
MV = the value of the mortgage to the borrower,
RFNC = cost of refinancing the mortgage,
$h_o = H_o - H$, where H_o is the value of the property to the owner, and H is the current market value of the property.

If the borrower is offered the opportunity to relocate, wealth depends on whether the borrower defaults or pays off the loan. If the borrower does not default, wealth is determined by nonhousing wealth, NW, book equity, E, selling costs, SC, the present value of the income effects of relocating, Y, and moving costs, MC:

$$W = NW + E - SC + Y - MC.$$

If the borrower defaults, wealth is determined by nonhousing wealth, NW, the present value of the income effects of relocating, Y, moving costs, MC, and the present value of the cost of default to the borrower, K:

$$W = NW - K + Y - MC.$$

[33] The 37 loans that defaulted would have dependent variables of zero (i.e., non-default) for every year except the year they defaulted. Therefore, the analysis includes only 37 observations that defaulted out of the 2,073 observations used in the analysis.

This implies that, if the borrower must relocate, default will be optimal if $-K > E - SC$. For default to be optimal, net equity must be negative and the amount of negative net equity must be larger than the present value of the cost of defaulting.

Giliberto and Houston (1989) show conditions under which the relocation decision is independent of the default decision and conditions under which the decisions are mutually dependent. Let δ represent the present value of relocating. Let α represent the incremental wealth change from selling the residence and β represent the incremental wealth change from default. A borrower will always choose to relocate if $\alpha + \delta > 0$. The authors refer to this as a relocating borrower. Default will be optimal for the relocating borrower if $\beta > \alpha$ ($-K > E - SC$).

Now consider a borrower facing a relocation opportunity where the present value of relocating is positive ($\delta > 0$), but not high enough to make up for negative equity in the property ($\alpha < 0$ and $\alpha + \delta < 0$). This borrower will relocate and default only if $\delta + \beta > 0$. Giliberto and Houston (1989) refer to this as a marginal relocator. This borrower will never relocate without defaulting.

Finally, consider a borrower facing a relocation opportunity where $\delta < 0$. This borrower will never sell the property and relocate. To do so would reduce the borrower's net wealth. Giliberto and Houston (1989) refer to this as a default relocator. The borrower will only relocate if the increase in wealth from defaulting is sufficiently high to make $\delta + \beta > 0$. The borrower will default if the net gain from transferring the property loss to the lender is greater than the net cost of relocating.

This model indicates that the decision to default may be influenced by the value of relocation opportunities. For the relocating borrower, the incremental wealth change from default could be negative and default will still be optimal as long as $\beta > \alpha$. For the default relocator, the incremental wealth change from default must be positive for default to be optimal since the present value of income from relocating is negative ($\delta < 0$). This indicates that the relocation decision and the default decision are independent under some circumstances and mutually dependent in other circumstances.

Another type of second generation research employs option-based models of default. In these models, default is viewed as a put option. Each payment period the borrower has the option to sell the home to the lender for the balance of the loan (Quercia and Stegman (1992)). In the option-based models, the borrower decides whether or not to exercise the default option based solely on the equity the borrower has in the property. If the borrower has negative equity, then default is expected.

Foster and Van Order (1984) develop two theoretical models for default. The first model is a pure option-based model where the borrower will default if the market value of the residence falls below the present value of the mortgage. In this model, the number of defaults, D_t, is equal to the number of loans with negative equity, C_t.

The second model assumes there are transactions costs associated with selling the residence and with exercising the option to default on the mortgage. It also assumes something happens to force the borrower to move out of the residence. Given that the borrower must move out of the residence, default will not occur if the equity in the home is greater than the costs associated with selling the residence.

Default may occur if net equity is negative. Let e represent net equity and d represent the transactions costs associated with defaulting on the mortgage. The borrower can default and pay d or sell the residence and receive e (or pay if e is negative). If net equity is negative such that e < -d, the borrower will be better off defaulting than selling the residence.

According to Foster and Van Order (1984), borrowers may also default when net equity is negative but exceeds the cost of exercising the default option (0 > e > -d). Assume P is the market price of the home and $\overline{P}$ is the borrower's subjective value of the home. Given negative net equity, P - $\overline{P}$ represents the borrower's subjective loss in holding the house. If the subjective loss from holding the house exceeds the net cost of default (e + d), the borrower will default. Foster and Van Order's (1984) second model is written as:

$$D_t = p_1 C_{1t} + p_2 C_{2t},$$

where D_t is the number of defaults on the FHA-insured loans in the sample for each time period t, C_{1t} is the number of loans in the sample where net equity is negative, but greater than the cost of exercising the default option, C_{2t} is the number of loans in the sample where net equity is negative and less than the cost of exercising the default option, and p_1 and p_2 represent the probability of default given C_1 and C_2, respectively. According to the model, all borrowers in C_2 should default, and many of the borrowers in C_1 should default, but default may not be immediate. Therefore, lagged values of C_1 and C_2 may be important. If default is immediate for borrowers in C_2, lagged values of C_2 will not be important and p_2 should be equal to one.

Foster and Van Order (1984) use FHA data from 1960 through 1978 that includes information from the endorsement year through the disposition of the loan. The authors calculate contemporaneous equity by simulating changes in house prices under the assumption that price changes follow a symmetric, stable Paretian distribution. They also calculate loan balances and the market value of the mortgages.[34] The final regression model includes C_2, C_1, and C_1 lagged one, two, and three periods, the expected inflation rate, the divorce rate, the housing expense to income ratio, and the age of the loan. All equity related variables enter the model through C_1 and C_2. Foster and Van Order (1984) use ordinary least squares (OLS) regression to estimate the model. The empirical tests lead Foster and Van Order (1984) to reject the simple option-based model of default that ignores transactions costs, but they find support for the model that incorporates transactions costs.

Riddiough (1991) focuses on using option-based methodology to price mortgages. Like Foster and Van Order (1984), he views the default decision as the intersection of two events. Riddiough (1991) hypothesizes that equity is the major determinant in the borrower's default decision, but borrowers with negative equity will continue paying their mortgage unless a trigger event occurs that forces them to move from their residence.

---

[34] Foster and Van Order's empirical tests rely on several assumptions. Loan balances are calculated assuming borrower's pay exactly the required payment each month and the required payment is calculated using average loan rates for the endorsement year. In their empirical tests, Foster and Van Order (1984) also assume the p's are constant across the sample.

C. Third Generation Research

Third generation research does little to add to the theoretical models of the borrower's default decision (Quercia and Stegman (1992)). These investigations focus on (1) expanding the discussion on the role of transactions costs in the option-based models of default and (2) estimating the fraction of large loan pools that will default.

Kau, Keenan, and Kim (1994) investigate the importance of transactions costs in the borrower's default decision. They develop a partial differential equation for the probability of default under the assumption that borrowers will make optimal decisions with regard to wealth. Based on the partial differential equation developed for the probability of default, the authors use numerical analysis to determine default probabilities for mortgages.

Using base parameters of 90 percent loan-to-value ratio, 9 percent contract interest rate, 8 percent market interest rate, 30 year mortgage, 10 percent interest rate volatility and 10 percent house price volatility, Kau, Keenan and Kim (1994) calculate the probability of default over the life of the loan to be 5.15 percent. Lowering the loan-to-value ratio to 80 percent decreases the probability of default to 0.76 percent while increasing the loan-to-value ratio to 95 percent increases the probability of default to 10.52 percent.

House price variances also play an important role in the model. If the variance in house prices increases from 10 percent to 15 percent, the probability of default increases from 5.15 percent to 11.63 percent. Decreasing the variance in house prices to 5 percent reduces the probability of default to 0.28%.

The authors provide an example using the base parameters to show that default will not be optimal as soon as net equity becomes negative, even in the absence of transactions costs. They use an example where a borrower takes out a $90,000 loan on a house valued at $100,000. Again, using the base parameter values provided earlier, the authors calculate the amount of house price depreciation that would be necessary to make default optimal after two years. They find that default will not be optimal after two years unless the market value of the house drops below $77,234. This represents a 23 percent decline in the value since loan origination (2 years) and the borrower's equity position in the house would be approximately -$12,200.[35]

This result is not explained by transactions costs, since there are no transactions costs in the model. Kau, Keenan and Kim (1994) argue that borrowers will not find it optimal to exercise the default option until equity is significantly negative because to do so requires they give up the right to exercise the option in the future. The default option has value, and that value is lost when the mortgage is terminated.[36]

Kau, Keenan and Kim (1994) conclude there is no evidence that transactions costs play a role in borrowers' decisions whether or not to exercise the default option. Based on their analysis and observed default rates in the market, the authors conclude that borrowers exercise the default option when it is rational to do so.

Lekkas *et. al.* (1993) test the option-based model of default to determine if the empirical evidence is consistent with the frictionless model or if transactions costs play a

---

[35] The authors calculate the principal balance of the loan after two years to be $89,436.
[36] The option to prepay the mortgage is also lost when the mortgage is terminated.

role in the default decision. Transactions costs that may influence the decision to exercise the default option include moving costs and reputation costs.

Lekkas *et. al.* (1993) develop four empirically testable propositions about loan loss severities that should hold if the frictionless model of default is accurate (i.e., if transactions costs don't matter). The first proposition is that the severity of losses on defaulted loans should fall as initial loan-to-value ratio increases. This proposition arises from the assumption that high loan-to-value ratio loans will have mortgage insurance. The effect of mortgage insurance premiums is to increase the effective coupon rate on the loan. Increasing the effective coupon rate on the loan increases the cost of keeping the option alive. Therefore, borrowers paying mortgage insurance premiums will find default to be optimal before borrowers not paying mortgage insurance premiums, all else equal. The result is that high loan-to-value ratio loans will have a higher probability of default and less loss severity on defaulted loans than low loan-to-value ratio loans.

The second proposition is that loss severity on defaulted loans should be independent of the region of the country or the year the loan is originated. The economic conditions in different regions will influence default rates, but should not influence loss severities on defaulted loans if the frictionless model holds.

The third proposition is that loss severity on defaulted loans should decrease with the age of the mortgage. As the mortgage ages, the time premium associated with the default option gets smaller, decreasing the value of the default option. Therefore, the older the mortgage is, the earlier the default option will be exercised.

The fourth proposition is that loss severity on defaulted loans should decrease as the differential between the coupon rate and the current market interest rate increases (coupon rate minus current market rate). Higher coupon rates increase the cost of keeping the option alive, so borrowers will exercise the option earlier.[37]

Lekkas *et. al.* (1993) use data on single-family, owner-occupied loans purchased by Freddie Mac during the period 1975-1990. Loss severity is measured as the difference in the mortgage balance and the value of the house at default. The value of the house at default is measured two ways: (1) the appraised value at the time of default, and (2) the sales price received by Freddie Mac when the home is sold.

The authors test the propositions by estimating a regression model with six dummy variables for different ranges of loan-to-value ratios, the age of the mortgage, the coupon minus current market rate, and a dummy variable for loans on property located in Texas. The dependent variable is the actual loss severity on defaulted loans. The model is estimated for both specifications of the value of the house at default.

Lekkas *et. al.* (1993) find loss severities increase with loan-to-value ratios, contrary to the first proposition. Loss severities are also significantly higher for Texas loans than for other loans, contrary to the second proposition. The age of the mortgage is found to have a negative impact on loss severities from default, consistent with the third proposition. Finally, the sign of coupon minus current rate is ambiguous in the two specifications of the dependent variable and is not significant in either specification. These findings lead the authors to reject the frictionless model of default.

---

[37] If the differential is significant enough, the borrower will want to terminate the mortgage, either through prepayment (refinancing) or default. This should make the borrower more likely to default at all levels of negative equity.

Hendershott and Schultz (1993) conduct an empirical investigation of the determinants of claims on FHA-insured single-family mortgages. Using FHA data on loans insured during the 1975-1987 period, the authors estimate a model to predict the conditional claim rate for each year. The conditional claim rate is the number of claims made during the policy year, t, on loans originated in year, y, that are still in the loan pool at the beginning of the policy year, t.

The data is subdivided into seven categories based on the size of the loan. Therefore, for any given origination year and policy year, there are 7 separate calculated conditional claim rates. For example, loans originated in 1975 are placed into one of seven size categories. For each policy year, 1975-1990, there are seven conditional claim rates for loans originated in 1975. Since there are 16 policy years (1975 through 1990), there are 112 separate conditional claim rates calculated on mortgages originated in 1975.[38]

Independent variables included in the model are (1) the expected percentage market equity, (2) the expected percentage book equity, (3) a measure of house price dispersion since the origination date, and (4) the unemployment rate. These variables are used to estimate a semi-log probability model as shown:

$$\ln CCR_{s,y,t} = \sum_{i=1}^{13} \alpha_i D_{i,t} + \sum_{i-1}^{7} \beta_j D_s (EM_{s,y,t} - \overline{EM}) + \beta_8 (EB_{s,y,t} - \overline{EB}) + \beta_9 (U_t - \overline{U}) + \beta_{10} (DISP_{y,t} - \overline{DISP}) + \varepsilon_{s,y,t},$$

[38] For loans originated in 1987, there are only four policy years. Therefore, there would be 28 conditional claim rates for loans originated in 1987. In all, there are 931 conditional claim rates.

where $CCR_{s,y,t}$ is the conditional claim rate, $D_{i,t}$ equals one in policy year i and zero otherwise. The percentage market equity, EM, percentage book equity, EB, unemployment rate, U, and home price dispersion, DISP, are measured as deviations from their mean values. The first term in the equation is the claim rate in policy year i for average economic conditions (i.e., when $EM = \overline{EM}$, $EB = \overline{EB}$, $V = \overline{V}$, $DISP = \overline{DISP}$).

Hendershott and Schultz (1993) find all four independent variables are important in explaining conditional claim rates. They find claim rates are negatively related to the percentage market equity and the percentage book equity. Higher claim rates are positively related to house price dispersion and unemployment rates. Claim rates increase rapidly as the initial loan-to-value ratio is increased. Loans with initial loan-to-value ratios of 96 percent are 20 times more likely to be foreclosed in the fifth policy year than loans with initial loan-to-value ratios of 80 percent. The authors also find the impact of market and book equity are three times higher for loans with initial loan-to-value ratios of 95-97 percent than for loans with initial loan-to-value ratios of 80 percent.

Schwartz and Torous (1993) use a methodology similar to Hendershott and Schultz (1993) to investigate mortgage prepayment and default. The authors estimate separate Poisson regression models to predict prepayments and defaults using Freddie Mac data on 1-4 family, 30-year fixed rate loans originated from 1975 to 1990. The dependent variable in the default equation is the quarterly probability of default.

The data is identified by origination year and policy year as in Hendershott and Schultz (1993). Explanatory variables in the default model include (1) dummy variables for the region of the country, (2) initial loan-to-value ratios, (3) housing return and volatility in return, (4) interest rate data including interest rate volatility and the slope of the term structure, and (5) the age of the mortgage.

The authors find that mortgages in the Southwest have the highest probability of default, and mortgages in the Northeast have the lowest probability of default. They also find the initial loan-to-value ratio and the volatility in housing returns are positively related to loan default. The degree to which increases in the initial loan-to-value ratios increase the probability of default is not the same for all regions. Increasing the initial loan-to-value ratio from 80 percent to 90 percent increases the quarterly probability of default from about 0.4 percent to 0.8 percent in the Southwest, but only increases the quarterly probability of default from 0.15 percent to 0.20 percent in the Northeast.[39] The impact of housing index volatility on the probability of default is also much greater for mortgages originated in the Southwest than for mortgages originated in the Northeast.

D. *Current State of the Research on Residential Loan Defaults*

Investigations of residential loan defaults consistently find home equity influences the default decision. Option-based models of default indicate home equity should have the dominant effect. The current debate on the option-based models of default focus on the role of transactions costs and the value of deferring the exercise of the default option.

Several investigations have provided evidence that other factors besides home equity influence the default decision. Vandell and Thibodeau (1985) and Giliberto and Houston (1989) develop models of the default decision where non-equity factors

[39] This assumes other variables are held constant at their mean values.

influence the default decision. They also indicate that default is more likely if some "trigger event" forces the borrower to move.

The model of the loan decision in the next chapter incorporates a model of the borrower's default decision. The model of the borrower's default decision reflects the importance of home equity and the role of a "trigger" or "crisis" event on the default decision.

CHAPTER IV.
THEORY AND HYPOTHESES

A. Introduction

This chapter develops a theoretical loan decision model and a theoretical analysis of discrimination in lending. The theoretical loan decision model is developed to show what factors are important in determining whether or not a loan should be approved. The theoretical lending model assumes owner wealth maximizing behavior by lenders.

The theoretical analysis of discrimination in lending develops expected relationships in loan denial and default rates that can result from discrimination when credit quality is identical for nonminority and minority applicants (homogeneous) and when nonminority applicants have a higher distribution of credit quality (tend to be more creditworthy) than minority applicants (heterogeneous). The analysis also shows the effect of noise in estimating credit quality when the distribution of credit quality is heterogeneous.

The owner wealth maximizing model developed in this chapter will be used to develop an equal outcome model in the next chapter. The owner wealth maximizing model and the equal outcome model should include all of the same variables if the distribution of credit quality is homogeneous with respect to race. If the distribution of credit quality is heterogeneous with respect to race, the equal outcome model will include some, but not all, of the variables in the owner wealth maximizing model.

B. *A Theoretical Mortgage Loan Decision Model*

This section develops a theoretical model of the credit decision which assumes all positive risk-adjusted net present value (NPV) loans will be accepted. The purpose of this model is to provide the theoretical basis for an owner wealth maximizing empirical model of the credit decision that does not consider race. The empirical owner wealth maximizing model will be estimated to determine if it produces racial disparities in mortgage lending decisions. A solid theoretical basis for the empirical model is important to establish that the empirical model correctly represents credit decisions that maximize shareholder wealth without consideration for race. Without this development, pure racial prejudice or a faulty credit decision process cannot be ruled out. This theoretical credit decision model assumes lenders accept all positive net present value (NPV) loans and borrowers make rational decisions to maximize the present value of their wealth. The theoretical model also provides a detailed analysis of the borrower's default decision and incorporates the default analysis into the credit decision model. Most previous models of the credit decision are empirical models with little or no theoretical justification. A few of the models, such as Munnell *et. al.* (1996) are based on maximizing the profitability of the loan by minimizing the probability of default. However, these models do not include any analysis of the borrower's default decision. They simply include empirical variables the investigator(s) think should measure default risk.

1. Assumptions of the Model

The theoretical mortgage lending model makes the following simplifying assumptions: (1) all loans are fixed-rate loans, (2) all loans have the same term to maturity, (3) all loans are fully amortizing loans requiring principal and interest payments at the beginning of each month, (4) all loans are either granted at the prevailing market interest rate or denied, and (5) prepayment risk exists but it is not considered in making the decision of whether or not a borrower meets creditworthiness standards to be approved for a loan. The assumption that all loans are fixed-rate loans simply implies the model developed in this chapter could not be directly applied to underwriting adjustable rate loans. Underwriting adjustable rate loans requires additional considerations and these loans are not the focus of this research. Empirical tests will use only fixed-rate loans.

The assumption that all loans have the same term to maturity implies that a lender underwriting loans to two different borrowers at the same time will apply the same term premium to both loans. Using this assumption, the model will develop the argument that default risk is the primary determinant in whether or not a loan is approved. All loans in the empirical tests will have the same loan term to maturity.

The assumption that all loans are fully amortizing loans requiring principal and interest payments at the beginning of each month along with the assumption that all loans have the same loan term to maturity leads to the conclusion that the timing of expected cash flows from all residential mortgage loans will be the same. All loans in the empirical tests are fully amortizing loans.

The assumption that all loans are either approved at the prevailing market mortgage rate simplifies the model and is consistent with previous residential mortgage lending models and current underwriting practices. The majority of the evidence suggests that credit is rationed in this manner in the single family residential mortgage market. Tootell (1993) and Munnell *et. al.* (1996) assume a single market interest rate for mortgages that is applied to all loans that meet minimum underwriting standards. Tootell (1993) cites King (1980) as providing evidence supporting the theory of a single market interest rate.

Lenders selling loans in the secondary market will be unlikely to charge lower rates for higher quality loans, since the loans would have to be sold at a lower price in the secondary market. Lenders will also be hesitant to charge higher rates to marginally less qualified minority borrowers and take the risk of being sued for discrimination. Barefoot (1995) indicates lenders are concerned about the "regulatory risks" of risk-based pricing even though it has been encouraged by Federal Reserve Board Chairman Alan Greenspan. Barefoot (1995) points to Justice Department settlements with American Family Mutual Insurance Company of Madison, Wisconsin and First National Bank of Vicksburg, Mississippi where risk-based pricing led to charges of racial discrimination. Barefoot (1995) also states that the Department of Housing and Urban Development (HUD) prohibits charging interest rates that vary by more than two points in the same geographic area.

The assumption that prepayment risk exists but does not play a role in the loan underwriter's decision also simplifies the model and also appears to be consistent with actual underwriting practices. Theoretical and empirical research on the determinants of mortgage prepayments indicate the dominant factors are the equity the borrower has in the home and the difference between the market interest rate and the contract interest rate on the loan. Prepayments are believed to be positively related to home equity and negatively related to the market interest rate minus the contract interest rate.[40] Research on loan defaults indicate the higher the borrower's home equity, the less likely the borrower is to default.[41] Since prepayment risk and default risk are inversely related, lenders face a tradeoff between high prepayment risk and high default risk. Therefore, the ramifications of prepayments and defaults must be considered to determine which will play a dominant role in the loan underwriting process.

The loss from prepayment depends on whether or not the loan is sold in the secondary market. The model makes no assumption in this regard. If the lender holds the loan in the portfolio, the maximum loss from prepayment is the opportunity cost of continuing to earn the contract rate of interest. No principal is lost. If the loan is sold in the secondary market, there is no loss from prepayment to the originator of the loan.

The loss from default also depends on whether or not the loan is sold in the secondary market. If the lender holds the loan in the portfolio, there are real losses from default. Even if the principal balance is eventually recovered, there are economic losses associated with time delays before the loan is settled. If the loan is sold in the secondary market without recourse, there is no immediate financial loss from the default but if

---

[40] Kau *et. al.* (1993) and Foster and Van Order (1985) hypothesize these relationships. Foster and Van Order (1985) test their model empirically and find strong support for these relationships.

[41] This relationship has been found in numerous investigations. Examples are von Furstenberg (1969), Jackson and Kaserman (1980) and Campbell and Dietrich (1983).

default rates on loans sold by a particular lender are perceived to be too high the lender may not be able to sell loans in the secondary market in the future.

It is obvious that lenders should prefer high prepayment risk to high default risk. Since there is an inverse relationship between prepayment risk and default risk, the model assumes lenders place more importance on default risk. Therefore, the model allows for borrower prepayment, but prepayment risk is not considered in making the decision of whether or not a borrower meets creditworthiness standards to be approved for a loan.

This appears to be consistent with actual underwriting practices. Lenders know borrowers may prepay loans but loan underwriters do not deny loans to borrowers who are high prepayment risks. This is also supported by empirical research. Research that models the loan decision of the lender and empirically tests these models find higher home equity (i.e., larger downpayments) increases the probability the loan will be approved. According to the expected influence of home equity on the probability of default and prepayment, this indicates lenders focus on the probability of default and do not deny loans to individuals who are high prepayment risks.

It follows from the inverse relationship between the probability of default and the probability of prepayment that borrowers with high prepayment risk will tend to have low default risk. Rejecting applicants for being high prepayment risks would eliminate the most creditworthy borrowers from the pool of borrowers. Lenders face considerable legal liability if creditworthy minority applicants are rejected for being high prepayment risks. Lenders would have considerable difficulty convincing regulators and the judicial system that an otherwise highly qualified minority applicant was denied a loan because the lender feared the applicant would pay off the loan prior to maturity.

2. *The Credit Decision of the Lender*

a. *The Net Present Value Rule, Value Additivity, and Shareholder Wealth Maximization*

One of the foundations of finance theory is that the primary goal of a firm's managers is the maximization of shareholder wealth. Theory indicates the NPV rule should be used in making investment decisions because this approach satisfies the value-additivity principle and maximizes shareholder wealth. The value additivity principle says the value of the firm is the sum of the value of the firm's investment projects. Assuming there are N projects, the value of the firm is:

$$V = \sum_{j=1}^{N} V_j \tag{1}$$

where V_j is the value of investment j. The value additivity principle holds that projects can be evaluated independently of one another.

The NPV of an investment is found by calculating the present value of its cash flows at the discount rate appropriate for the risk of the investment and then subtracting the required initial investment (i.e., the cost of the investment). A positive NPV indicates the value of the investment is greater than the cost of the investment. The aggregate wealth of the firm will increase by the NPV of the investment project. Managers should undertake all positive NPV investments to maximize shareholder wealth. Therefore, the correct decision rule for evaluating investment proposals is to accept all proposals that have a NPV greater than zero. This decision rule will maximize shareholder wealth.

b. The Net Present Value Rule and Financial Institutions

Managers of financial institutions seek investments that will maximize shareholder wealth. The loan selection process can be viewed as an investment decision where the manager of the financial institution chooses only those investments with positive NPV's because this will maximize shareholder wealth.

The decision to grant a fixed-rate residential mortgage loan can be based on the NPV rule. The NPV is determined by (1) the amount and timing of the cash flows and (2) the required rate of return, r_i, on the loan. The amount of the cash flows from a fixed-rate residential mortgage loan is determined by the dollar amount of funds loaned and the contract mortgage interest rate.

The contractual rate of interest, r, is a function of (1) the risk-free rate of interest, r_f, (2) a term premium, T_p, and (3) a default risk premium, $\overline{D_p}$, which is the default risk premium for marginal loans.[42] The default risk premium is the same for all loans that meet or exceed the underwriting standards because it was argued that it is not economically rational for mortgage lenders to charge different interest rates on mortgage contracts with the same terms.[43] There is no default risk pricing in the loan terms. The contractual rate is defined as:

---

[42] The default risk premium, $\overline{D_P}$, on marginal loans is the default risk premium that would apply to a borrower with qualifications identical to the minimum underwriting standards that will permit sale on the secondary market. This assumes lenders want to have the option to sell loans in the secondary market.
[43] This assumes all loans have the same term to maturity. The empirical research uses a sample of 30 year fixed-rate mortgage loans.

$$r = r_f + T_p + \overline{D}_p \tag{2}$$

The borrower's required rate of return, r_i, on the loan is determined by (1) the risk-free rate of interest, r_f, (2) a term premium, T_p, and (3) a default premium, D_p:

$$r_i = r_f + T_p + D_p \tag{3}$$

If two borrowers apply for 30 year fixed rate mortgage loans at the same time, any differences in the lender's required rate of return, r_i, on the loans must be due to the difference in the default premiums since the risk-free rate and the term premium are identical for the two borrowers. The default premium, D_p, is unique to each loan and is determined by the probability the borrower will default on the loan and the amount of the expected loss from default.

The NPV of a mortgage loan is a function of the stream of cash flows (loan payments), the required rate of return, and the loan amount, where

$$\text{LOAN PAYMENT} = \frac{\text{LOAN AMOUNT}}{\frac{1}{r} - \frac{1}{r(1+r)^t}}, \tag{4}$$

where r = coupon interest rate on the loan (prevailing market mortgage interest rate), and

$$\text{NPV} = \sum_{t=1}^{T} \frac{\text{LOAN PAYMENT}_t}{(1+r_i)^t} - \text{LOAN AMOUNT}, \tag{5}$$

where r_i = the required rate of return on the loan. Whether or not a loan's NPV is positive depends only on the relationship of the required rate of return on the loan, r_i, to the prevailing market mortgage interest rate, r. This relationship is directly dependent on the relationship of the default premium on the marginally qualified borrower, $\overline{D_p}$, and the specific default premium, D_p. If $D_p < \overline{D_p}$, then $r_i < r$ and the loan will have a positive

NPV. If $D_p > \overline{D_p}$, then $r_i > r$ and the loan will have a negative NPV. The NPV will decrease as r_i increases:

$$\frac{\partial \text{NPV}}{\partial r_i} < 0 \tag{6}$$

The risk free rate, term premium and the default premium for the loan are embedded in the prevailing market mortgage interest rate. The required rate of return, r_i, will be greater than the prevailing market interest rate, r, if the borrower's default premium is higher than the default premium on the marginal loan, i.e., $D_p > \overline{D_p}$. The marginal loan will have a net present value of zero. The determination of whether the NPV on a loan will be positive or negative depends primarily on the borrower's probability of default and the expected loss from default.

The magnitude of the NPV is also influenced by the amount of the loan. Given a value of $r_i < r$, the NPV will be higher for larger loans:

$$\frac{\partial \text{NPV}}{\partial \text{LOANAMT}} > 0 \tag{7}$$

Given a value of $r_i > r$, the NPV will be lower (more negative) for larger loans:

$$\frac{\partial \text{NPV}}{\partial \text{LOANAMT}} < 0. \tag{8}$$

Since the lender denies loans when $r_i > r$, only the case where $r_i < r$ is relevant.
In the absence of capital rationing, all loans with a required rate of return, r_i, below the prevailing market interest rate, r, will be approved. The only borrower-specific factor used to determine the required rate of return on the loan is the default risk premium for the borrower, which is a measure of the lender's assessment of the probability of default

and the expected loss from default. The higher the probability of default and expected loss from default, the higher the default risk premium for the borrower. The default risk premium for marginal loans, $\overline{D_p}$, reflects the maximum acceptable probability of loss and expected loss from default. Therefore, the probability of default and expected loss from default are the major determinants in the loan decision. The probability of loan approval, P(APPROVE), is a function of (1) the probability the borrower will default on the loan, P(DEFAULT), and (2) the amount of the expected loss from default, EXPLOSS:

$$\text{P(APPROVE)} = \text{f(P(DEFAULT), EXPLOSS)}. \tag{9}$$

The probability the loan will be approved is inversely related to the probability of default and the amount of the expected loss from default:

$$\frac{\partial \text{P(APPROVE)}}{\partial \text{P(DEFAULT)}}, \frac{\partial \text{P(APPROVE)}}{\partial \text{EXPLOSS}} < 0. \tag{10}$$

The borrower's default decision must be incorporated into the credit decision framework because the probability of default is the major determinant in the credit decision. The model of the borrower's default decision is presented in the next section.

3. *The Default Decision From the Borrower's Perspective*

Borrowers are assumed to maximize utility by maximizing wealth. Therefore, they are assumed to prefer more wealth to less wealth, and to make rational decisions consistent with maximizing the present value of their expected future cash flows. This is consistent with the NPV rule (Brealey and Myers (1988)). Borrowers make decisions regarding loan repayment at the beginning of each month, t. A flowchart of the borrower's default decision involves four decision points and is shown in Figure 5.

The first decision point (#1) involves an examination of net equity. The borrower will not default if the following condition holds:

$$V_t - C - L_t > 0,$$

Where V_t equals the current market value of the residence, C equals the costs associated with selling the residence and L_t is the principal balance outstanding on the loan. If the borrower's net equity is positive, default will not occur because it would (at least) cost the borrower the positive equity (V_t - C - L_t)[44]. At this point, if a borrower experiences a crisis event that forces the borrower to move from the residence, the borrower is better off selling the residence than defaulting on the loan (since V_t - C - L_t is positive). The second decision point (#2) shows that even if the borrower's equity is negative, default will not occur if the value the borrower places on a default-free credit rating, REP, exceeds the negative net equity to be gained from defaulting (REP > L_t - (V_t- C)).[45] If a crisis event occurs in this case, the gain from defaulting would not be worth the loss of a default-free credit rating.

If the net equity is negative and greater than the value of a default-free credit rating (L_t - (V_t - C) > REP), then the borrower is better off to default than to sell. In this case, either of two conditions are sufficient for default to be optimal. If, for some

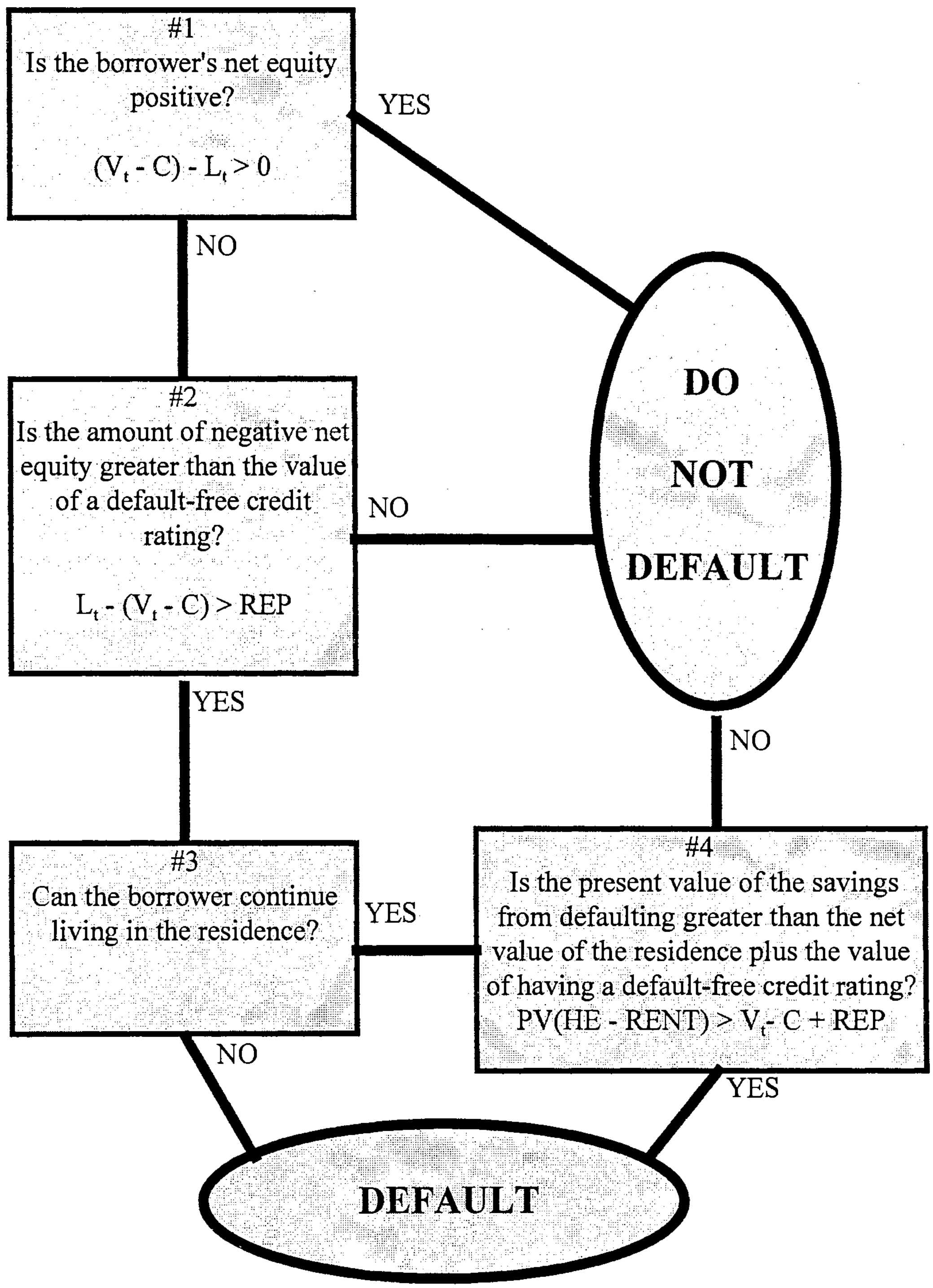
---

[44] This is rational and is consistent with option-based models of default such as Foster and Van Order (1984).

[45] Lekkas *et. al.* (1993) argue that reputation costs may prevent borrowers from defaulting when the default option is in the money.

FIGURE 5

THE BORROWER'S DEFAULT DECISION



reason, the borrower must move from the residence, default is optimal (decision point #3). The inability to continue living in the residence could result from a disruption in employment, a significant reduction in income, a significant increase in expenses, or some other crisis event. These events are compounded if the borrower lacks sufficient financial resources to overcome the crisis event.

A second condition sufficient to justify default may exist if the present value of the savings from defaulting is greater than the net value of the residence plus the value of a default-free credit rating even if the borrower can continue living in the residence (decision point #4). The borrower can continue to pay the monthly cost of homeownership, HE, which includes principal, interest, taxes, insurance, maintenance costs, less the tax benefits from homeownership. Continuing payments entitles the borrower to maintain ownership of the residence, which has a net value of $V_t - C$. The borrower will also be able to maintain a default-free credit rating, REP. Therefore, the benefit of continuing to make payments is $V_t - C + REP$. The benefit from defaulting on the loan is the present value of the difference between the monthly cost of homeownership and the monthly cost to rent a comparable residence, RENT. Default will only be optimal if the present value of the savings from defaulting is greater than the net value of the residence plus the value of a default-free credit rating ($PV\{HE - RENT\} > V_t - C + REP$). The discount rate used to calculate the present value of the savings from defaulting is the borrower's cost of nonhousing debt. If the present value of the savings from defaulting is greater than the net value of the residence plus the value of a default-free credit rating, the homeowner could obtain funds at the nonhousing debt rate of r* and use the monthly savings from defaulting to repay the loan.

To summarize, default will be optimal when the borrower's equity is negative and the benefit from defaulting exceeds the value of a default-free credit rating, $L_t - \{(V_t - C\} > REP)$, if one of two conditions exists. First, default will be optimal if the borrower cannot continue living in the residence. Second, default will be optimal if the present value of the savings from defaulting is greater than the net value of the residence plus the value of a default-free credit rating ($PV\{HE - RENT\} > V_t - C + REP$).

The borrower's default decision is therefore a function of (1) the market value of the residence, V_t, (2) costs associated with selling the residence, C, (3) the outstanding loan balance, L_t, (4) the value the borrower places on having a default-free credit rating, REP, (5) the monthly cost of homeownership, HE, (6) the monthly rental cost for a comparable dwelling, RENT, (7) the borrower's cost of nonhousing debt, r*, (8) the financial resources of the borrower, FINRES, and (9) the probability of a crisis event, P(CRISIS):

$$P(DEFAULT) = f(V_t, C, L_t, REP, HE\ RENT, r^*, FINRES, P(CRISIS)). \qquad (11)$$

Increases in the current market value of the residence, V_t, increase the net equity of the borrower and reduce the probability of default. Increases in the value placed on having a default free credit rating, REP, decrease the probability of default. Higher values of REP mean negative net equity must be higher for default to be optimal. Increases in the cost to rent a comparable dwelling, RENT, make default less likely, since the savings from defaulting will be smaller. Increases in the borrower's cost of nonhousing debt, r*, decrease the present value of the benefits from defaulting and decrease the probability of default. Increases in financial resources, FINRES, decrease the probability of default because borrowers with strong financial resources are more likely to be able to overcome other negative factors.

Increases in the costs associated with selling the residence, C, and the outstanding loan balance, L_t, reduce the net equity of the borrower and increase the probability of default. Increases in the monthly cost of homeownership, HE, increase the savings from defaulting, making default more likely. Finally, the higher the probability of a crisis event, P(CRISIS), the higher the probability of default. Crisis events may force borrowers to move, making default more likely. The marginal impact of each variable on the probability of default is:

$$\frac{\partial P(DEFAULT)}{\partial V_t}, \frac{\partial P(DEFAULT)}{\partial REP}, \frac{\partial P(DEFAULT)}{\partial RENT}, \frac{\partial P(DEFAULT)}{\partial r^*}, \frac{\partial P(DEFAULT)}{\partial FINRES} < 0,$$

$$\frac{\partial P(DEFAULT)}{\partial C}, \frac{\partial P(DEFAULT)}{\partial L_t}, \frac{\partial P(DEFAULT)}{\partial HE}, \frac{\partial P(DEFAULT)}{\partial P(CRISIS)} > 0 \qquad (12)$$

4. *The Full Mortgage Loan Decision Model*

The lender's decision to approve a loan is hypothesized to be a function of the probability of default and the amount of the expected loss from default. The previous section developed a framework for the borrower's default decision. The theoretical credit decision model for fixed-rate residential loans is completed by incorporating the borrower's default decision into the loan approval decision. If equation 11 is substituted into equation 9, the decision to approve a loan is a function (1) the value of the residence, V_t, (2) the costs associated with selling the residence, C, (3) the outstanding loan balance,

L_t, (4) the value the borrower places on having a default-free credit rating, REP, (5) the monthly cost of homeownership, HE, (6) the monthly rental rate for a comparable dwelling, RENT, (7) the borrower's cost of nonhousing debt, r*, (8) the financial resources of the borrower, FINRES, (9) the probability of a crisis event, P(CRISIS), and (10) the expected loss from default, EXPLOSS:

$$P(APPROVE) = f(V_t, C, L_t, REP, HE, RENT, r^*, FINRES, P(CRISIS), EXPLOSS). \qquad (13)$$

Any factor that increases the probability of default decreases the probability of loan approval. The probability of default increases with increases in the costs associated with selling the residence, C, the outstanding loan balance, L_t, the monthly cost of homeownership, HE, and the probability of a crisis event, P(CRISIS). Therefore, increases in these factors decrease the probability of loan approval. Increases in the value of the residence, V_t, the value the borrower places on having a default-free credit rating, REP, the monthly rental cost for a comparable dwelling, RENT, the borrower's cost of nonhousing debt, r*, and the financial resources of the borrower, FINRES, decrease the probability of loan default. These variables should be positively related to the probability of loan approval. The marginal impact of each variable on the probability that the loan will be approved is shown:

$$\frac{\partial P(APPROVE)}{\partial C}, \frac{\partial P(APPROVE)}{\partial L_t}, \frac{\partial P(APPROVE)}{\partial HE}, \frac{\partial P(APPROVE)}{\partial P(CRISIS)} < 0, \qquad (14)$$

$$\frac{\partial P(APPROVE)}{\partial V_t}, \frac{\partial P(APPROVE)}{\partial REP}, \frac{\partial P(APPROVE)}{\partial RENT}, \frac{\partial P(APPROVE)}{\partial r^*}, \frac{\partial P(APPROVE)}{\partial FINRES} > 0.$$

The theoretical credit decision model hypothesizes the only borrower-specific risk that is considered is the risk of default. Therefore, default risk is believed to play the dominant role in the credit decision. The theoretical credit decision model developed in this chapter is estimated with an empirical model, referred to as the owner wealth maximizing model.

C. *A Theoretical Analysis of Discrimination in Lending*

As developed in Chapter I, racial disparities in lending can result from three things: (1) taste-based discrimination, (2) owner wealth maximizing lending decisions, and (3) statistical discrimination. This investigation assumes that any taste-based discrimination in lending takes the form of marginal minority applicants being held to a higher standard than nonminority applicants. Longhofer (1996) indicates that taste-based discrimination occurs by requiring members of the "disfavored" group to meet higher cutoff standards. Hunter and Walker (1995) find race plays no role in the outcome of loan applications when applicants are of high quality, but race significantly impacts outcomes for marginal loan applications. This indicates that taste-based discrimination either does not exist, or it occurs only among marginally qualified minority applicants. Statistical discrimination occurs when lenders use variables correlated with race that do not measure the probability of loan default or the expected loss from default. Lenders use proxies to determine the creditworthiness of the borrower. Therefore, the creditworthiness of the borrower is measured with noise. This research assumes that the majority of all discrimination is statistical discrimination, and that discrimination occurs only among marginally qualified applicants.

The impact of statistical discrimination on loan denial and default rates for nonminority and minority borrowers depends on two factors. The first factor is whether the distribution of credit quality is the same for all borrowers (i.e., homogeneous), or whether the distribution of credit quality is heterogeneous with respect to race. The second factor is the amount and form of noise in the estimation of borrower creditworthiness.

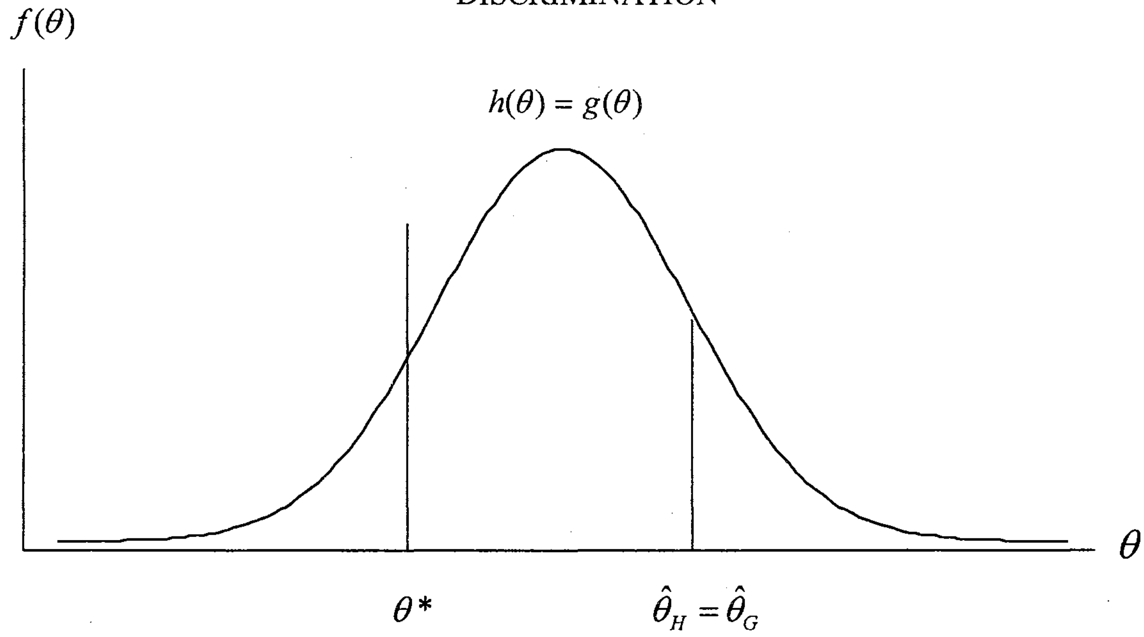
1. Homogeneous Borrowing Population With No Noise

Ferguson and Peters (1995) show that the relationship between relative denial and default rates for nonminority and minority borrowers in a nondiscriminatory lending environment depend critically on their relative distributions of credit quality. Ferguson and Peters (1995) develop a model where the credit score of an applicant, θ, represents the probability of repayment. The probability of default for each borrower, by definition, is $1 - \theta$. A uniform, nondiscriminatory lending policy requires that all loans with credit quality above some cutoff, θ^*, are approved and all loans with credit quality below θ^* are denied.

If nonminority and minority borrowers have the same distribution of credit quality, the borrowing population is homogeneous. An illustration of the Ferguson and Peters (1995) model with no discrimination and homogeneous credit quality is shown in Figure 6. Since the distribution of credit quality is identical, the average credit quality of approved minority borrowers, $\hat{\theta}_H$, is equal to the average credit quality of approved nonminority borrowers, $\hat{\theta}_G$. With homogeneous credit quality, nonminority and

FIGURE 6

HOMOGENEOUS CREDIT QUALITY WITH NO DISCRIMINATION



minority borrowers should experience equal denial and default rates under a uniform, nondiscriminatory lending policy. If minorities are held to a higher standard, $\theta^* + \delta$, they will have higher denial rates and the average credit quality of approved minority borrowers will be greater than the average credit quality of approved nonminority borrowers. This is shown in Figure 7. Since the average default rate for minority borrowers is $1 - \hat{\theta}_H$ and the average default rate for nonminority borrowers is $1 - \hat{\theta}_G$, minority borrowers will have lower default rates than nonminority borrowers in a discriminatory lending environment if credit quality is homogeneous.[46]

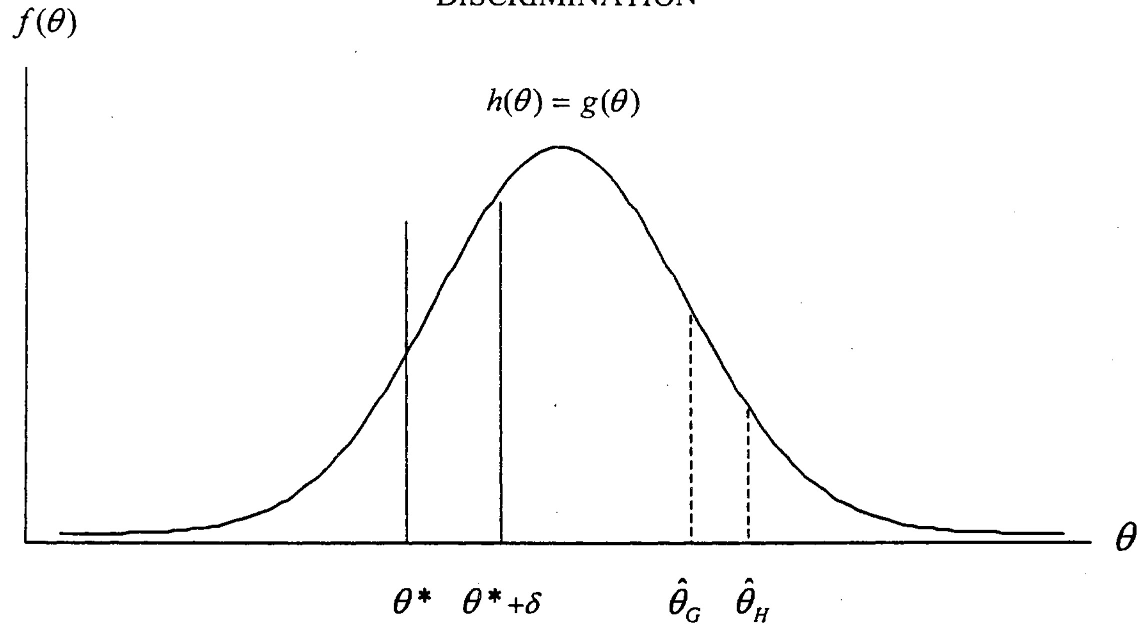
2. *Heterogeneous Borrowing Population With No Noise*

Ferguson and Peters (1995) point out that recent empirical evidence indicates the distribution of credit quality is heterogeneous with respect to nonminority and minority borrowers. Munnell *et. al.* (1996) find minority borrowers have, on average, greater debt burdens, higher loan-to-value ratios, and weaker credit histories than nonminority borrowers. Ferguson and Peters (1995) develop a model where the distribution of credit quality for nonminority borrowers first-order stochastically dominates the distribution of credit quality for minority borrowers. Under this condition, Ferguson and Peters (1995) show that nonminority borrowers will have higher average credit quality than minority

[46] This is the argument used by Becker (1993) and others to refute the findings of Munnell *et. al.* (1996). An implicit assumption of this argument is that the distribution of credit quality is homogeneous with respect to race.

FIGURE 7

HOMOGENEOUS CREDIT QUALITY WITH DISCRIMINATION



borrowers (i.e., $\hat{\theta}_G > \hat{\theta}_H$).[47] This is shown in Figure 8. Since the average credit quality for nonminority borrowers, $\hat{\theta}_G$, is greater than the average credit quality for minority borrowers, $\hat{\theta}_H$, minority borrowers will have higher denial and default rates than nonminority borrowers under a uniform, nondiscriminatory lending policy.[48] From Figure 8, it is obvious that minority borrowers will have higher denial rates than nonminority borrowers under a uniform credit policy where all applicants with creditworthiness above θ^* are approved and all applicants with creditworthiness below θ^* are denied.[49] Also, because $G(\theta)$ first-order stochastically dominates $H(\theta)$ and $\hat{\theta}_H < \hat{\theta}_G$, minority borrowers will have higher default rates than nonminority borrowers.[50]

---

[47] Since the lending policy approves all loans with credit quality above θ^*, marginal minority applicants and marginal nonminority applicants may have identical characteristics but the average credit quality will be higher for nonminority borrowers than for minority borrowers.

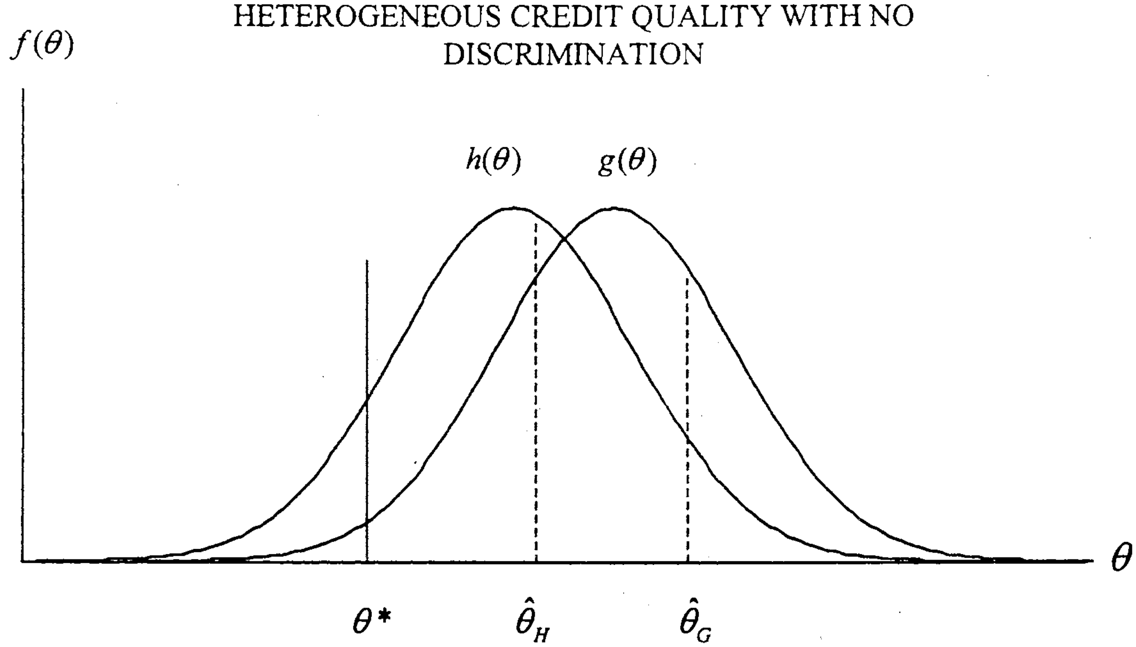
[48] The explanation of racial disparities caused by owner wealth maximizing loan decisions assumes credit quality is heterogeneous with respect to race. If credit quality is homogeneous with respect to race, credit rationing will not affect minorities disproportionately.

[49] The area under the curve to the left of θ^* represents denied applications. The area under the probability density function, $h(\theta)$, to the left of θ^* is greater than the area under the probability density function, $g(\theta)$ to the left of θ^* ($\int_0^{\theta^*} h(\theta)d\theta > \int_0^{\theta^*} g(\theta)d\theta$).

[50] The average default rate for minority borrowers is $1 - \hat{\theta}_h$ and the average default rate for nonminority borrowers is $1 - \hat{\theta}_g$.

FIGURE 8

HETEROGENEOUS CREDIT QUALITY WITH NO DISCRIMINATION



3. *Implications of a Heterogeneous Borrowing Population*

The Ferguson and Peters (1995) investigation has important implications for research on discrimination in lending. One implication is that, assuming the borrowing population is heterogeneous, it is very difficult to discern discrimination from prudent lending by looking at relative denial and default rates. Ferguson and Peters (1995) show that only two combinations of denial and default rates lead to a finding of discrimination when the borrowing population is heterogeneous. First, if nonminority borrowers have higher denial rates and lower default rates than minority borrowers, there is evidence of reverse discrimination. Second, if minority borrowers have higher denial rates and lower default rates than nonminority borrowers, there is evidence of discrimination against minority borrowers. The expected relationship is that minority borrowers will have higher denial and default rates than nonminority borrowers in a nondiscriminatory lending environment when the distribution of credit quality is heterogeneous.[51]

4. *The Effect of Noise in the Estimation of θ*

As Ferguson and Peters (1995) show, the relative distribution of credit quality for nonminority and minority borrowers is an important consideration when examining racial disparities in residential mortgage lending. An equally important consideration is the

[51] Minority borrowers could also have higher denial and default rates than nonminority borrowers when there is discrimination against minority borrowers or discrimination against nonminority borrowers. Ferguson and Peters (1995) show that, when the borrowing population has heterogeneous credit quality, the combination of higher denial and default rates for minority borrowers is consistent with no discrimination, discrimination against minorities, and discrimination against nonminority.

effect of noise in estimating the credit quality of potential borrowers. Let θ represent the borrower's true probability of repayment. Lenders must estimate θ based on information they collect in underwriting the loan, since the borrower's true θ is unobservable to the lender. To the extent that variables used by the lender are imperfect proxies for the true probability of repayment, there is noise in the estimation of θ. Noise is defined here as incorrect estimates of the true probability of repayment, θ, resulting from the use of imperfect or incorrect variables in the calculated probability of repayment.

Let θ_c represent the lender's calculation of the probability of repayment for the borrower. The true probability of repayment, θ, is equal to the calculated probability of repayment, θ_c, plus a measurement error, ε: $\theta_w = \theta_{c_w} + \varepsilon_w$ for nonminority borrowers, and $\theta_b = \theta_{c_b} + \varepsilon_b$ for minority borrowers. Since the true probabilities of repayment, θ_w and θ_b, are unknown, the measurement errors, ε_w and ε_b, are unobservable.

Ferguson and Peters (1997) show that the effect of measurement errors on racial disparities in lending depend on (1) whether the distribution of credit quality is homogeneous or heterogeneous with respect to race, and (2) whether or not the errors are correlated with race. Assume measurement errors are uncorrelated with race and the distribution of credit quality is homogeneous with respect to race.[52] The measurement errors should have the same impact on minority and nonminority borrowers.

---

[52] Ferguson and Peters (1997) assume that a correlation between underwriting errors and race implies that errors are made over a wider range of θ's for minority borrowers than for nonminority borrowers. They assume the underwriting errors are symmetrical. Therefore, if errors are correlated with respect to race, Ferguson and Peters (1997) assume $\overline{\varepsilon_b} = \overline{\varepsilon_w}$ and $\sigma_{\varepsilon_b} > \sigma_{\varepsilon_w}$.

Now assume the measurement errors are uncorrelated with race, but the distribution of credit quality is heterogeneous with respect to race. Ferguson and Peters (1997) show that the measurement errors will have a more adverse impact on minority applicants than on nonminority applicants. That is, more creditworthy minority applicants than creditworthy nonminority applicants will be denied loans. Their analysis assumes that the majority of nonminority applicants and the majority of minority applicants meet the minimum requirements for loan approval (i.e. $\bar{\theta} > \hat{\theta}$). The result is that the marginal credit score is on the upward sloping side of the distribution for both nonminority and minority borrowers (Figure 9). Therefore, if underwriting errors occur in a random manner (and are symmetrical as Ferguson and Peters assume), the errors will result in a higher proportion of creditworthy applicants being denied mortgages than the proportion of uncreditworthy applicants that are approved.[53] If the distribution of minority borrowers is lower than the distribution of nonminority borrowers (i.e., nonminority borrowers have a higher distribution of credit quality), then minorities will be more adversely affected by the underwriting errors than nonminority borrowers.[54]

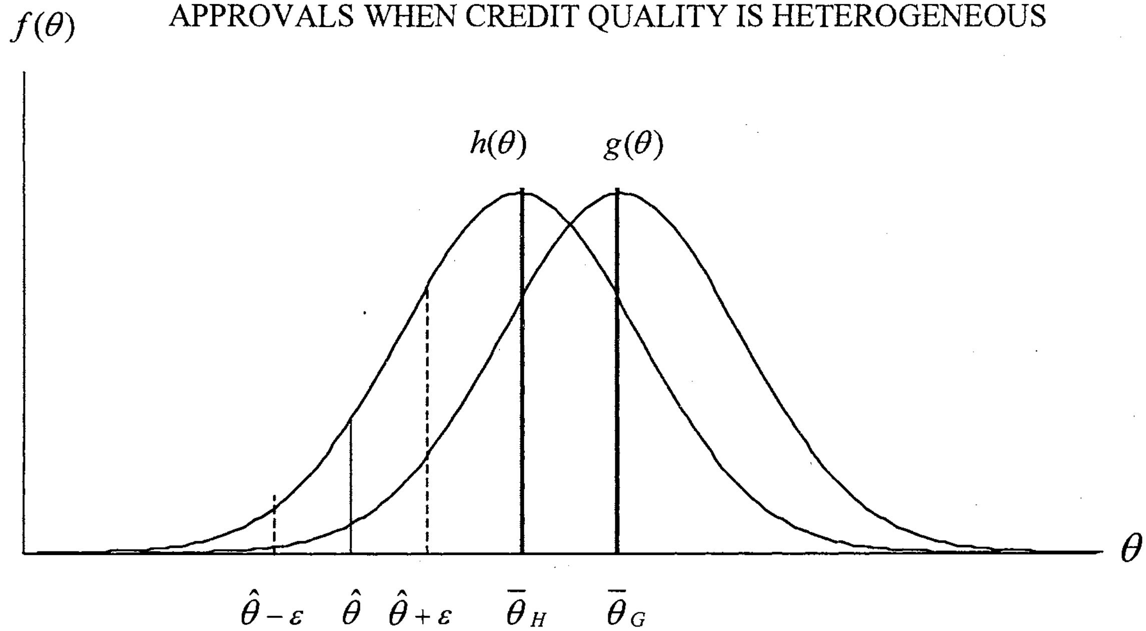
---

[53] The number of creditworthy minorities denied loans will be greater than the number of uncreditworthy minorities granted loans. Assuming p is the proportion of underwriting errors, the number of uncreditworthy minorities granted loans is $\int_{\hat{\theta}-\varepsilon}^{\hat{\theta}} ph(\theta)d\theta$. This is less than the number of creditworthy minorities denied loans, $\int_{\hat{\theta}}^{\hat{\theta}+\varepsilon} ph(\theta)d\theta$.

[54] This is true because $\int_{\hat{\theta}}^{\hat{\theta}+\varepsilon} ph(\theta)d\theta - \int_{\hat{\theta}-\varepsilon}^{\hat{\theta}} ph(\theta)d\theta > \int_{\hat{\theta}}^{\hat{\theta}+\varepsilon} pg(\theta)d\theta - \int_{\hat{\theta}-\varepsilon}^{\hat{\theta}} pg(\theta)d\theta$.

FIGURE 9

THE IMPACT OF RANDOM UNDERWRITING ERRORS ON LOAN APPROVALS WHEN CREDIT QUALITY IS HETEROGENEOUS



Next, assume the measurement errors are correlated with race, such that lenders make more underwriting errors when reviewing applications from minority borrowers. Ferguson and Peters (1997) show that more underwriting errors will increase unwarranted racial disparities in lending decisions regardless of the distribution of credit quality if underwriting errors are symmetric and the majority of all loan applications are approved.[55]

The next three sections expand the Ferguson and Peters (1997) analysis with examples. The first section shows the impact of using an owner wealth maximizing lending model when the distribution of credit quality is heterogeneous. The second section shows the impact of using a variable in the model that is a poor proxy for true creditworthiness if (1) the variable is uncorrelated with race, or (2) the variable is positively correlated with race and believed to be negatively correlated with creditworthiness. The third section shows the impact of using an equal outcome model to make lending decisions if the actual distribution of credit quality is heterogeneous. An equal outcome model is a subset of an original model constructed to assure equal outcomes in lending decisions.

a. The Impact of Using an Owner Wealth Maximizing Model

Assume the probability of repayment, θ, is determined by factors A, B, and C that are unobservable to the lender ($\theta = f(A, B, C)$), where A and B are uncorrelated with race

[55] The assumption that the majority of all loans are approved implies that the cutoff point for marginal loans is on the upward sloping side of the distribution. If errors are symmetric, this means there will be more good loans rejected than bad loans approved.

and C is correlated with race. The distribution of θ will be heterogeneous since factor C is correlated with race. Assume the lender's calculation of the probability of repayment, θ_c, is determined by factors W, X, and Y ($\theta_c = f(W, X, Y)$), where W, X, and Y are the best available proxies for factors A, B, and C. The calculation of θ_c will contain noise, but the noise is minimized by using the best proxies available. Assuming W and X are uncorrelated with race and Y's correlation with race is approximately equal to C's correlation with race, the distribution of θ_c will be approximately equal to the distribution of θ.[56] In this case, approving all loans with calculated credit quality above θ^* will be consistent with maximizing owner wealth because θ_c is approximately equal to θ across the borrowing population. Assuming the majority of all loan applications are approved and credit quality is heterogeneous, any noise will have a disproportionate impact on minority borrowers.[57]

b. The Impact of Using Poor Proxies for Creditworthiness

Now assume the calculation of θ_c includes variable Z in addition to variables W, X, and Y ($\theta_c = f(W, X, Y, Z)$), where variable Z is not a good proxy for any of the factors, A, B, and C.[58] Since variable Z is used in the determination of θ_c and is not a good proxy for any of the determinants of θ, including variable Z increases the noise in θ_c. Including

[56] The only way the distributions would be equal is if variables W, X, and Y were perfect proxies for factors A, B, and C.
[57] This is based on the assumptions in the previous section that (1) credit quality is heterogeneous, and (2) the majority of all loan applications are approved.
[58] This does not assume that lenders know Z is not a good proxy. Lenders may believe Z is a good proxy and use it in underwriting loans when, in fact, it is a very poor proxy.

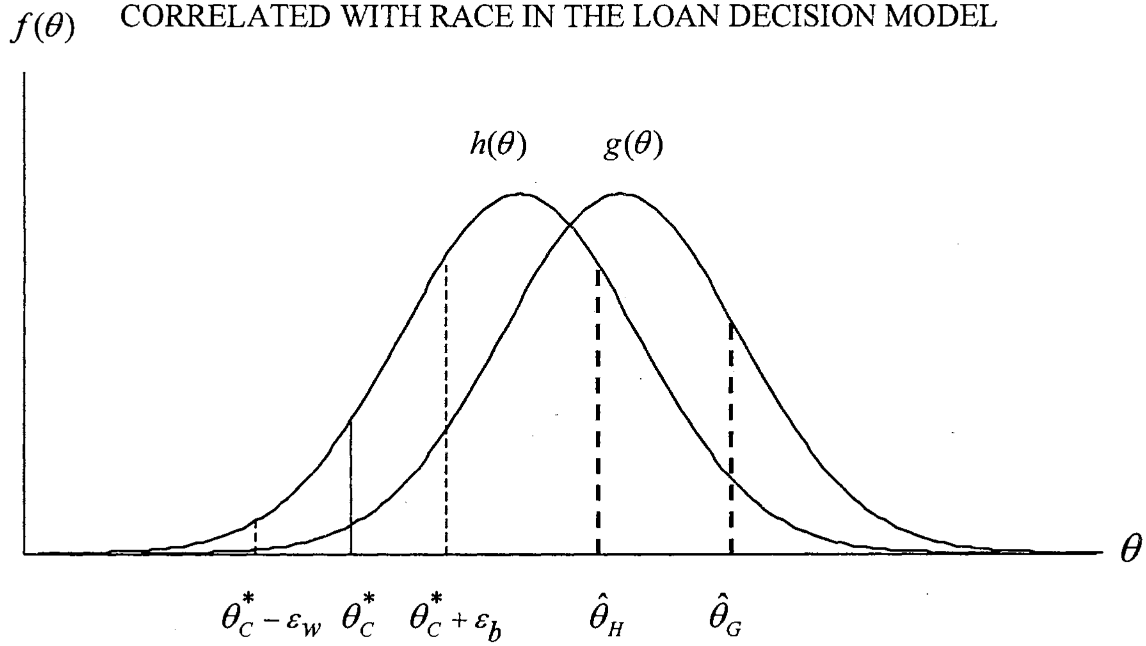
variable Z results in a faulty credit decision model that either approves some loans that should be rejected, rejects some loans that should be approved, or both. Even if Z is uncorrelated with race, the noise in θ_c will have a more detrimental impact on minority borrowers than on nonminority borrowers if credit quality is heterogeneous. This is not considered discrimination because the errors caused by including variable Z should be identically distributed for marginal nonminority and minority borrowers (since Z is uncorrelated with race). If Z is correlated with race such that θ_c is systematically lower than θ for minority borrowers and θ_c is systematically higher than θ for nonminority borrowers, the result is statistical discrimination. In this case, the errors that result from including variable Z tend to understate the credit quality of minority borrowers and overstate the credit quality of nonminority borrowers. This is shown in Figure 10. The amount of the understatement of θ for minority borrowers is ε_b ($\theta_c = \theta - \varepsilon_b$), and the amount of the overstatement of θ for nonminority borrowers is ε_w ($\theta_c = \theta + \varepsilon_w$). The credit decision model approves all loans where $\theta_c > \theta_c$*.[59] Marginal borrowers will have $\theta_c = \theta_c$*. Therefore, the true credit quality of marginal minority borrowers will be θ_c* + ε_b and the true credit quality of marginal nonminority borrowers will be θ_c* - ε_w. The marginal nonminority borrower will have credit quality $\varepsilon_b + \varepsilon_w$ lower than the marginal

---

[59] Ferguson and Peters (1995 and 1997) assume the marginal cutoff point is known. An example of the cutoff point could be all loans that meet the minimum underwriting guidelines for sale in the secondary market are approved and all others are rejected. Note that the underwriting guidelines do not measure true creditworthiness. They use proxies to measure creditworthiness. Therefore, the marginal cutoff point used by lenders reflects calculated creditworthiness, not true creditworthiness.

FIGURE 10

THE IMPACT OF USING POOR PROXIES THAT ARE CORRELATED WITH RACE IN THE LOAN DECISION MODEL



minority borrower. Therefore, discriminatory lending occurs due to the use of variable Z in the credit decision model. This is statistical discrimination. This illustrates the importance of selecting variables to be included in a lending model carefully, because errors in variable selection may increase unwarranted racial disparities in lending. If the variables used are correlated with race and are poor proxies of creditworthiness, statistical discrimination will occur.

c. The Impact of Using an Equal Outcome Model

If θ_c is heterogeneous, we expect to see differences in loan denial and default rates for nonminority and minority borrowers. One way to eliminate the racial disparity in loan denials is to use a credit decision model that will produce a distribution of θ_c's that is homogeneous.[60] Let the true probability of repayment, θ, be determined by factors A, B, and C as before. Factors A and B are uncorrelated with race and factor C is correlated with race. Let θ_c be determined by variables W and X, which are the best available proxies for factors A and B. Assuming variables W and X are uncorrelated with race, the distribution of θ_c will be homogeneous, but the distribution of θ remains heterogeneous. There will be an increase in the noise in θ_c using this model in place of the owner wealth maximizing model.

Assume the correlation of factor C to race is such that minority borrowers tend to have higher values of factor C and higher values of factor C are associated with lower

---

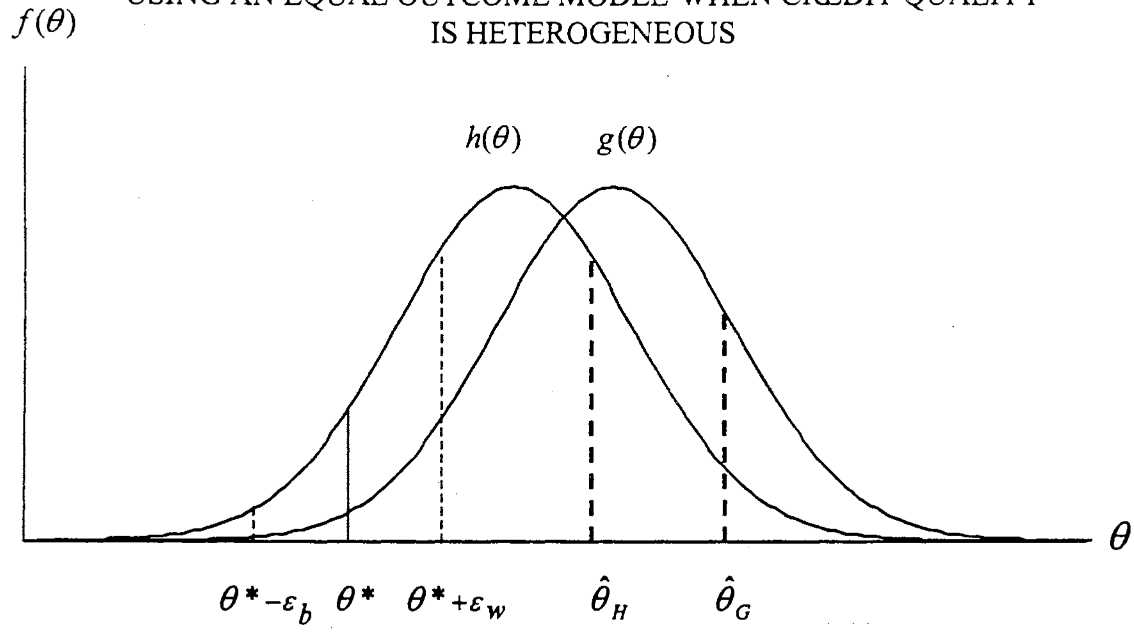
[60] This should eliminate racial disparities in loan denials. Any model that reduces the magnitude of heterogeneity will reduce racial disparities in loan denials.

probabilities of repayment. The omission of variables to proxy factor C will result in a tendency to overstate the probability of repayment for minority borrowers and understate the probability of repayment for nonminority borrowers. Assume omitting variables that proxy factor C systematically overstates θ for minority borrowers by ε_b ($\theta_c = \theta + \varepsilon_b$) and systematically understates θ for nonminority borrowers by ε_w ($\theta_c = \theta - \varepsilon_w$). This is shown in Figure 11. Marginal borrowers will have $\theta_c = \theta^*$. The true credit quality of marginal minority borrowers will be $\theta^* - \varepsilon_b$ and the true credit quality of marginal nonminority borrowers will be $\theta^* + \varepsilon_w$. Marginal nonminority borrowers will have credit quality $\varepsilon_b + \varepsilon_w$ higher than marginal minority borrowers. Average credit quality for minority borrowers, $\hat{\theta}_H$, will be significantly less than the average credit quality for nonminority borrowers, $\hat{\theta}_G$. The equal outcome model will achieve equal loan approval rates for nonminority and minority borrowers, but default rates should be significantly higher for minority borrowers.

There is more noise in the estimation of θ using the equal outcome model than using the owner wealth maximizing model, because the equal outcome model suffers from an omitted variable bias. Ferguson and Peters (1997) show that increasing the noise increases the variability in the measurement error. Given some true probability of repayment, θ', the owner wealth maximizing model will calculate the probability of repayment, θ'_c, to be $\theta' + \varepsilon_1$, while the equal outcome model will calculate θ'_c to be $\theta' + \varepsilon_2$. Since the owner wealth maximizing model contains less noise than the equal outcome

FIGURE 11

USING AN EQUAL OUTCOME MODEL WHEN CREDIT QUALITY IS HETEROGENEOUS



model, there should be less dispersion in the measurement errors using the owner wealth maximizing model than using the equal outcome model ($\sigma_{\varepsilon_1} < \sigma_{\varepsilon_2}$). This is shown in Figure 12. Since there is less dispersion in the owner wealth maximizing model in estimating θ, the owner wealth maximizing model should outperform the equal outcome model in predicting loan defaults.

D. *Theoretical Hypotheses*

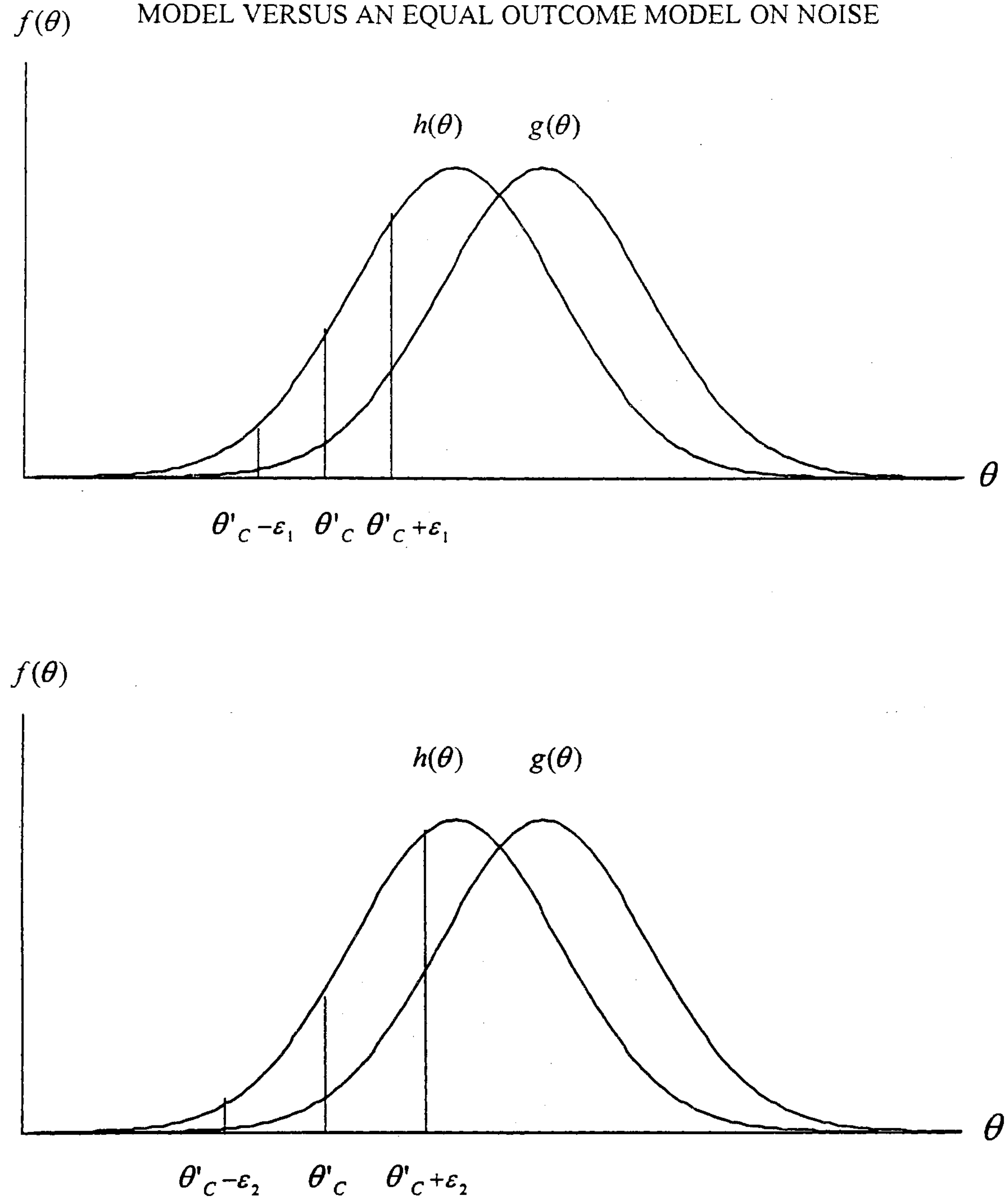
1. *Introduction*

The recent focus of policymakers on eliminating racial disparities in lending decisions leads to the question of what the costs and benefits of such actions would be. In order to eliminate racial disparities in lending decisions, it is necessary to eliminate the heterogeneity in credit decision models. If the true distribution of credit quality is homogeneous, equal outcome lending will have no cost. However, if the true distribution of credit quality is heterogeneous, the cost of equal outcome lending will be any loss in the ability of the equal outcome model to predict loan defaults compared to the owner wealth maximizing model.

Measuring the costs and benefits of using an equal outcome model instead of an owner wealth maximizing model is the focus of the second and third hypotheses. The benefit of using an equal outcome model is any reduction in the racial disparity in lending decisions from using the equal outcome model instead of the owner wealth maximizing model. This is tested in hypothesis two. The cost of using the equal outcome model is

FIGURE 12

THE EFFECT OF USING AN OWNER WEALTH MAXIMIZING MODEL VERSUS AN EQUAL OUTCOME MODEL ON NOISE



any reduction in default prediction from using the equal outcome model instead of the owner wealth maximizing model. The third hypothesis involves measuring this cost.

2. *Hypothesis I: The Distribution of Credit Quality*

An important consideration when investigating racial disparities in lending is whether the distribution of credit quality is heterogeneous or homogeneous with respect to race. Ferguson and Peters (1995) suggest that homogeneous distributions of the probability of repayment should result in equal denial and default rates for nonminority and minority borrowers in the absence of discrimination. This assumes the probability of repayment, θ, is measured without noise. It will also be true if any noise in estimating θ is symmetrical and uncorrelated with race. If the distribution of credit quality is heterogeneous, the relationship between denial rates, default rates and discrimination becomes more complex as shown in Ferguson and Peters (1997) and earlier in this research.

Previous empirical evidence indicates different denial and default rates occur for nonminority and minority applicants. These results are consistent with heterogeneous distributions of the probability of repayment, but do not rule out the existence of discrimination. Before conducting an analysis of whether or not discrimination exists in the market for residential mortgage loans, it is important to establish whether the distribution of θ is homogeneous or heterogeneous with respect to race. A finding of homogeneous distributions would indicate that the owner wealth maximizing model and the equal outcome model should be identical and that performance-based measures used by policymakers to determine if lenders discriminate will be effective. A finding of heterogeneous distributions indicates that the equal outcome model will not contain all of the same variables as the owner wealth maximizing model and that equal outcome lending encouraged by the performance-based measures will probably have a cost. The expectation is that the distributions of the probability of repayment for nonminority and minority borrowers are heterogeneous:

H_O: Minority borrowers are, on average, at least as creditworthy as nonminority borrowers.

H_A: Minority borrowers are, on average, less creditworthy than nonminority borrowers.

3. Hypothesis II: Racial Disparities in Lending Decisions

Empirical evidence indicates that minority applicants have higher denial rates than nonminority applicants. This may be due to heterogeneous credit quality, discrimination, or both. Differences in loan approval rates for nonminority and minority applicants will depend on the variables used to make the loan decision. As discussed earlier, one way to try to reduce or eliminate racial disparities in loan approvals is to use an equal outcome model to make loan decisions. An equal outcome model should result in less racial disparity in loan approvals than an owner wealth maximizing model:

H_O: The equal outcome model results in at least as much racial disparity in loan approvals as the owner wealth maximizing model.

H_A: The equal outcome model results in less racial disparity in loan approvals than the owner wealth maximizing model.

4. *Hypothesis III: Default Prediction*

The use of an equal outcome model may reduce racial disparities in lending but it will also increase the noise in the estimation of the probability of repayment, θ. Due to the increased noise, the measurement errors will be greater for the equal outcome model than for the owner wealth maximizing model. Since the equal outcome model is expected to give less accurate estimates of the probability of repayment, it should not perform as well in predicting loan defaults as the owner wealth maximizing model:

> H_O: The equal outcome model predicts loan defaults at least as well as the owner wealth maximizing model.

> H_A: The owner wealth maximizing model predicts loan defaults with more accuracy than the equal outcome model.

E. *Conclusion*

This chapter has developed a theoretical credit decision model that incorporates the borrower's default decision. The theoretical model is based on the premise that lenders make decisions that maximize shareholder wealth. The owner wealth maximizing model should minimize noise in calculating the probability of repayment. If variables in the owner wealth maximizing model are correlated with race, the distribution of the calculated probability of repayment, θ_c, will be heterogeneous. The first hypothesis involves testing the owner wealth maximizing model to determine if the borrowing population is heterogeneous.

The recent focus of policymakers on eliminating racial disparities in lending decisions leads to the question of what the costs and benefits of such actions would be.

In order to eliminate racial disparities in lending decisions, it is necessary to eliminate the heterogeneity in credit decision models. If the true distribution of credit quality is homogeneous, equal outcome lending will have no cost. However, if the true distribution of credit quality is heterogeneous, the cost of equal outcome lending will be any loss in the ability of the equal outcome model to predict loan defaults compared to the owner wealth maximizing model.

Measuring the costs and benefits of using an equal outcome model instead of an owner wealth maximizing model is the focus of the second and third hypotheses. The benefit of using an equal outcome model is any reduction in the racial disparity in lending decisions from using the equal outcome model instead of the owner wealth maximizing model. This is tested in hypothesis two. The cost of using the equal outcome model is any reduction in default prediction from using the equal outcome model instead of the owner wealth maximizing model. The third hypothesis involves measuring this cost. The next chapter develops the empirical analysis employed in testing the hypotheses.

CHAPTER V.
EMPIRICAL METHODOLOGY

A. Introduction

In their efforts to end discrimination in the mortgage market, policymakers in Washington have shifted the focus from evaluating lenders' efforts to lend to minorities to evaluating their performance in lending to minorities. Some industry experts caution that this type of assessment will lead to lending quotas or equal outcome lending, regardless of qualifications (England (1993)).

Equal outcome lending will accomplish the social goal of eliminating racial disparities in lending. The cost of this type of lending program to society depends on whether the racial disparities that exist are due primarily to discrimination or whether they occur because white applicants tend to be more creditworthy than minority applicants. The three hypotheses developed in the previous chapter address these issues.

The first hypothesis involves testing whether the distribution of credit quality is homogeneous or heterogeneous with respect to race. If the distribution of credit quality is homogeneous, the empirical owner wealth maximizing model developed in this chapter will produce equal loan approval rates across racial groups. If credit quality is heterogeneous with respect to race, the owner wealth maximizing model will produce racial disparities in loan approval rates. An equal outcome model is derived by testing all possible subsets of the empirical owner wealth maximizing model to determine which set of variables will produce equal loan approval rates for white and black borrowers. If the distribution of credit quality is heterogeneous with respect to race, the equal outcome model will not include all of the same variables that are in the owner wealth maximizing model.

The second hypothesis tests to determine if racial disparities in loan approvals produced by the equal outcome model are significantly less than the racial disparities produced by the owner wealth maximizing model. If the distribution of credit quality is found to be heterogeneous, the equal outcome model should result in significantly less racial disparity in loan approval rates than the owner wealth maximizing model. This hypothesis is tested using loan application data.

The third hypothesis tests the owner wealth maximizing model and the equal outcome model to determine if the owner wealth maximizing model is significantly better than the equal outcome model in predicting loan defaults. If the distribution of credit quality is heterogeneous with respect to race, the owner wealth maximizing model should significantly outperform the equal outcome model in loan default prediction. This hypothesis is tested using loan default data. This chapter describes the two datasets used in the empirical analysis, discusses the development of the empirical owner wealth maximizing model and equal outcome model, and outlines the statistical methodology used to test the hypotheses.

B. Data

One unique aspect of this research is the use of separate datasets to measure (1) racial disparity in lending decisions and (2) default prediction. The loan application data is important in measuring the racial disparity in lending decisions that will be produced by the owner wealth maximizing empirical model and the equal outcome model. The loan default data is important in testing each model's ability to predict loan defaults.

1. *Loan Application Data*

The loan application data used in this research is a partial dataset from Munnell, Tootell, Browne and McEneaney (1996), which contains information on loan applications taken by Boston area lenders in 1990. The dataset was released by the Boston Federal Reserve and includes most of the variables used by Munnell *et. al.* (1996), but does not include variables relating to the lender or the characteristics of the census tract. The loan application data is augmented with 1990 Census data from the U.S. Bureau of the Census.

Coding errors in the Munnell *et. al.* (1996) dataset have been detailed by Horne (1994), Liebowitz (1993), and Carr and Megbolugbe (1993). Carr and Megbolugbe (1993) subjected the Munnell *et. al.* (1996) dataset to several filters to "clean" the dataset. A similar process was performed on the data for this research. Observations that failed any of the following criteria were deleted from the sample:

Criterion #1: The purchase price must not exceed $10 million.
Criterion #2: Liquid assets must not exceed $10 million.
Criterion #3: The loan-to-value ratio must not exceed 1.2.
Criterion #4: The loan term must be between 120 and 500 months.
Criterion #5: The expense-to-income ratio must be less than the total obligations-to income ratio.

The initial sample consisted of 2,932 loan applications, 2,247 from white applicants, 471 from black applicants, and 214 from Hispanic (or other) applicants. The "cleaning" procedure reduced the initial sample of 2,932 loan applications to 2,447 loan applications, 1,914 from white applicants, 376 from black applicants, and 157 from Hispanic (or other) applicants. This research develops a model for fixed-rate single-family owner occupied mortgage loans. Only loans that meet these criteria are included in the final sample. All variable rate loans, loans secured by 2-4 residences, and loans secured by residences that are not owner-occupied are deleted from the final sample. Hispanic applicants are deleted from the dataset to make the loan application dataset consistent with the loan default dataset, which contains no Hispanic borrowers. The final sample consists of 1,255 loan applications, 1,096 from white applicants, 159 from black applicants. The final sample consists of 1,159 approved loans and 96 denied loans.

2. *Loan Default Data*

The loan default data used in this research consists of borrower, loan and property characteristics on FHA-insured single-family residential mortgage loans originated from 1986 through 1989. The data was recently made available by the U.S. Department of Housing and Urban Development (HUD). The dataset includes information on each FHA-insured loan from origination to the disposition of the loan, and indicates if the loan defaulted. The original dataset contains 104 variables and more than three million observations. The loan default data is augmented with 1990 Census data from the U.S. Bureau of the Census. The empirical tests of this research uses seven empirical variables derived from the theoretical model and a stratified random sample of 5,000 observations. The sample is stratified to assure an equal number of defaulted and non-defaulted loans in the sample.

Several restrictions are placed on the data used in the empirical tests. The sample data includes only those observations where the property is owner-occupied and the loan has a fixed-rate until maturity. All observations with missing values for any of the empirical variables are also excluded from the sample. The final sample of 5,000 observations includes 2,500 loans that defaulted and 2,500 loans that did not default. There are 4,000 white borrowers and 1,000 black borrowers in the sample, with forty-six states represented in the sample.

C. *Empirical Models*

1. *Owner Wealth Maximizing Model*

Implementation of the theoretical model requires translation of the theoretical variables into empirical variables. The proxy variables should be variables that can be observed by the lender at the time of the credit decision. The theoretical variables that must be estimated with available data are (1) the current market value of the residence, V_t, (2) the costs associated with selling the residence, C, (3) the outstanding loan balance, L_t, (4) the value the borrower places on having a default-free credit rating, REP, (5) the monthly cost of renting a comparable dwelling, RENT, (6) the borrower's cost of nonhousing capital, r*, (7) the financial resources of the borrower, FINRES, (8) the probability of a crisis event, P(CRISIS), and (9) the amount of the expected loss if default occurs, EXPLOSS:

$$P(APPROVE) = f(V_t, L_t, C, REP, HE, RENT, r^*, FINRES, P(CRISIS), EXPLOSS) \quad (15)$$

a. *The Value of the Residence, V_t, and the Outstanding Loan Balance, L_t*

The importance of the value of the residence and the outstanding loan balance is that they measure home equity. Therefore, these two variables should be measured relative to each other. The value of the residence is measured as the lower of the appraised value or the purchase price which is a more conservative approach than has been used in previous research. Previous research uses the appraised value to measure the value of the residence, without regard for the purchase price of the property. At loan origination, the outstanding loan balance is the loan amount. In this research, the variable used to proxy the relationship between the value of the residence and the outstanding loan balance is the ratio of the loan amount to the lower of the appraised value or purchase price:

$$LTV = \frac{LOAN\ AMOUNT}{\mathrm{MIN}(APPRAISED\ VALUE,\ PURCHASE\ PRICE)}. \tag{16}$$

Holding other factors constant, increases in the loan amount relative to the value of the residence will increase the probability of default and decrease the probability that the loan will be approved.

b. Costs Associated With the Sale of the Residence, C

The empirical model does not include proxies for the costs associated with the sale of the residence. The costs associated with selling the residence are expected to be proportional to the value of the residence. However, including the appraised value as an additional explanatory variable is not necessary since it would be collinear with the loan-to-value ratio. The real impact of these costs is whether or not they make net equity negative. If the loan-to-value ratio is sufficiently high, selling costs will make net equity negative. Selling costs will not make net equity negative if the loan-to-value ratio is low. Accurate information on selling costs is not available for the data used in the empirical analysis. The omission of this variable may cause the model to have an omitted variable bias.

c. The Value The Borrower Places on Having a Default-Free Credit Rating, REP

The value the borrower places on having a default free credit rating is measured with a dummy variable, CREDHIS, which equals 1 if the lender considered the applicant's credit history a positive compensating factor in the loan decision and 0 otherwise. The CREDHIS variable should be negatively related to the probability of default and positively related to the probability the loan will be approved.

For the loan application data, a second dummy variable is employed. The second dummy variable, BADCRED, equals 1 if the applicant has any history of loan defaults and 0 otherwise.[61] The BADCRED variable should be negatively related to the probability of loan approval.

d. Monthly Cost of Homeownership, HE, and Monthly Rental Costs, RENT

The monthly rental costs for a comparable dwelling should be measured relative to the monthly cost of homeownership. The monthly rental cost of a comparable dwelling is measured as the average monthly rent in the neighborhood. The monthly mortgage payment is used as a proxy for the monthly costs of homeownership. This research uses the ratio of the monthly mortgage payment to the average rental rate to

[61] The BADCRED variable is not available for the loan default dataset, therefore it cannot be tested to determine if it explains defaults.

measure the relationship between monthly rental costs and the monthly cost of homeownership, MTGRENT:

$$MTGRENT = \frac{MONTHLY\ MORTGAGE\ PAYMENT}{AVERAGE\ MONTHLY\ NEIGHBORHOOD\ RENT}. \tag{17}$$

Higher values of MTGRENT indicate the borrower can obtain alternative and comparable housing in the rental market at a much lower cost. The MTGRENT variable should be positively related to the probability of default and negatively related to the probability the loan will be approved. The monthly cost of homeownership should also include taxes, insurance, and maintenance costs, but that information is not available for the data used in the empirical analysis. The omission of the variables from the monthly cost of homeownership may result in the model having an omitted variable bias.

*e. The Borrower's Cost of Nonhousing Debt, r**

The empirical model does not include proxies for the borrower's cost of nonhousing debt. The borrower's cost of nonhousing debt cannot be obtained for the individual borrowers in the sample. The omission of this variable may result in the model having an omitted variable bias.

f. The Financial Resources of the Borrower, FINRES

The borrower's financial resources are important for two reasons. First, a borrower with strong financial resources may be able to overcome crisis events that would usually require a borrower to move from the residence. The ability to overcome crisis events will reduce the probability of default. Second, a borrower that places a high value on having a default-free credit rating still must have financial resources to protect against default. It is not enough that a borrower doesn't want to default. If a crisis event forces the borrower to move when the borrower has substantial negative home equity and very low financial resources, default may be the only possible course of action. In the same situation, a borrower with high financial resources might be able to sell the house and pay off the loan using financial resources to cover the negative equity. Borrowers with stronger financial resources should be less likely to default and more likely to have their loans approved.

Two empirical variables are used to measure the financial resources of the borrower: (1) the ratio of liquid assets to the mortgage payment, LIQMTG, and (2) the ratio of the borrower's residual income to the mortgage payment, RESMTG. Liquid assets and residual income are measured relative to the mortgage payment, because borrowers with lower mortgage payments are likely to be able to keep the loan current with lower levels of residual income and liquid assets. Higher ratios of LIQMTG and RESMTG should increase the probability of loan approval and decrease the probability of loan default.

g. The Probability of a Crisis Event, P(CRISIS)

Anything that increases the probability of a crisis event will increase the probability of default and decrease the probability that the loan will be approved. Crisis events are defined as any events that make it impossible for the borrower to continue living in the residence, e.g., loss of employment, significant reduction in income, and divorce. One empirical variable is used to proxy the probability of a crisis event. The variable is a measure of the job stability of the borrower. The empirical variable, EMPLOY, equals 1 if the lender considered the applicant's employment history a positive compensating factor in the loan decision and 0 otherwise. This variable should be positively related to the probability of loan approval and negatively related to the probability of loan default.

h. The Expected Loss From Default, EXPLOSS

If the borrower defaults on the loan, the lender may be exposed to a loss. The amount of loss exposure depends on the value of the collateral which is impacted by demand and supply factors for comparable dwellings in the neighborhood. The demand and supply of comparable dwellings in the neighborhood will be measured by the vacancy rate in the neighborhood. A dummy variable, VACANT, which equals 1 if the neighborhood vacancy rate exceeds five percent and 0 otherwise is used. The five percent threshold is used because vacancy rates of five percent or less are considered normal neighborhood vacancy rates. The dataset released by the Department of Housing and Urban Development (HUD) includes this dummy variable. Vacancy rates higher than five percent may be indicative of declining property values due to declining demand for housing in the neighborhood.

i. The Final Model

The empirical model developed from the theoretical model is shown:

$$
\begin{aligned}
P(APPROVE) = b_0 + b_1 LTV + b_2 CREDHIS + b_3 BADCRED + b_4 MTGRENT \\
+ b_5 LIQMTG + b_6 RESMTG + b_7 EMPLOY + b_8 VACANT + e
\end{aligned}
\tag{18}
$$

where LTV is the ratio of the loan amount to the lower of the purchase price or the appraised value, CREDHIS is a dummy variable that equals 1 if the applicant's credit history was a positive compensating factor in the loan decision and 0 otherwise, BADCRED is a dummy variable that equals 1 if the borrower has a history of loan defaults and 0 otherwise, MTGRENT is the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood, LIQMTG is the ratio of the borrower's liquid assets to the mortgage payment, RESMTG is the ratio of the applicant's residual income to the mortgage payment, EMPLOY is a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise, and VACANT is a dummy variable that equals 1 if the neighborhood vacancy rate exceeds five percent and 0 otherwise.[62]

2. Equal Outcome Model

The empirical equal outcome model is derived by testing all possible subsets of the empirical owner wealth maximizing model to determine which set of variables will produce equal average probabilities of loan approval for white and black applicants. This

---

[62] The empirical model used to test the loan default dataset does not include the BADCRED variable.

is achieved by comparing the average probability of loan approval for white and black applicants for all 127 possible subsets of the owner wealth maximizing model. The model that produces the closest average probability of loan approval for white and black applicants is the equal outcome model. The model is shown:

$$P(APPROVE) = b_0 + b_1 RESMTG + b_2 MTGRENT + b_3 EMPLOY + e \qquad (19)$$

The actual loan approval rates and the average probabilities of loan approval for white and black applicants using the two empirical models are shown in Table I. The average probability of loan approval for white and black applicants are approximately equal using the equal outcome model.

D. *Descriptive Statistics on the Data*

1. *Loan Application Data*

Descriptive statistics on the final sample of loan applications are shown in Table II. White applicants, on average, have significantly lower loan to value ratios (p-value $< .0001$) and significantly higher liquid asset to mortgage payment ratios than black applicants (p-value $< .0001$). The average loan to value ratio is 79.8 percent for white applicants compared to 86.5 percent for black applicants. White applicants in the sample are also more likely to have strong credit histories (p-value $< .0001$) and stable employment histories (p-value $< .0001$) than black applicants. The applicant's credit history is considered a positive compensating factor for 18.9 percent of white applicants and only 6.3 percent of black applicants. Employment history is a positive compensating factor for 31.1 percent of white applicants and 18.2 percent of black applicants.

TABLE I

AVERAGE PROBABILITIES OF LOAN APPROVAL FOR WHITE AND BLACK APPLICANTS

	White Applicants	Black Applicants
Actual Results: Approved Loans	93.80%	82.39%
Owner Wealth Maximizing Model	92.96%	88.17%
Equal Outcome Model	92.35%	92.36%

TABLE II

DESCRIPTIVE STATISTICS ON THE LOAN APPLICATION DATASET

Variable	Mean	Standard Deviation	Minimum	Maximum
All applicants (n=1255)				
Loan to value ratio	.806	.119	.500	1.188
Credit history is positive factor	.173	.378	.000	1.000
Previous loan default (BADCRED)	.124	.330	.000	1.000
Mortgage payment to average rent	.718	.578	.043	5.900
Liquid assets to mortgage payment	49.022	85.114	.000	1201.330
Residual income to mortgage payment	3.047	1.848	.113	22.345
Employment is positive factor	.295	.456	.000	1.000
Neighborhood vacancy > 5%	.306	.461	.000	1.000
White applicants (n=1096)				
Loan to value ratio	.798	.120	.500	1.188
Credit history is positive factor	.189	.392	.000	1.000
Previous loan default (BADCRED)	.105	.307	.000	1.000
Mortgage payment to average rent	.726	.599	.043	5.900
Liquid assets to mortgage payment	52.087	89.935	.000	1201.330
Residual income to mortgage payment	3.096	1.885	.113	22.345
Employment is positive factor	.311	.463	.000	1.000
Neighborhood vacancy > 5%	.281	.450	.000	1.000
Black applicants (n=159)				
Loan to value ratio	.865	.088	.521	1.071
Credit history is positive factor	.063	.244	.000	1.000
Previous loan default (BADCRED)	.258	.439	.000	1.000
Mortgage payment to average rent	.663	.403	.141	2.866
Liquid assets to mortgage payment	27.898	30.444	.000	195.666
Residual income to mortgage payment	2.709	1.534	.548	17.642
Employment is positive factor	.182	.387	.000	1.000
Neighborhood vacancy rate > 5%	.478	.501	.000	1.000

TABLE II (continued)

DESCRIPTIVE STATISTICS ON THE LOAN APPLICATION DATASET

Variable	Mean	Standard Deviation	Minimum	Maximum
Approved Applicants n=1159)				
Loan to value ratio	.803	.120	.500	1.188
Credit history is positive factor	.183	.387	.000	1.000
Previous loan default (BADCRED)	.107	.309	.000	1.000
Mortgage payment to average rent	.711	.567	.043	5.900
Liquid assets to mortgage payment	50.343	87.920	.000	1201.330
Residual income to mortgage payment	3.053	1.853	.113	22.345
Employment is positive factor	.296	.457	.000	1.000
Neighborhood vacancy > 5%	.297	.457	.000	1.000
Rejected Applicants (n=96)				
Loan to value ratio	.844	.093	.602	1.094
Credit history is positive factor	.052	.223	.000	1.000
Previous loan default (BADCRED)	.333	.474	.000	1.000
Mortgage payment to average rent	.800	.696	.164	4.704
Liquid assets to mortgage payment	33.079	33.513	.000	233.302
Residual income to mortgage payment	2.974	1.787	.834	12.160
Employment is positive factor	.281	.452	.000	1.000
Neighborhood vacancy > 5%	.417	.496	.000	1.000

TABLE II (continued)

DESCRIPTIVE STATISTICS ON THE LOAN APPLICATION DATASET

Tests for Difference in Population Means for White Vs. Black Applicants

Variable	White Applicants	Black Applicants	t-Statistic	p-Value
Loan to value ratio	.798	.865	- 8.52	<.0001
Credit history is positive factor	.189	.063	5.55	<.0001
Previous loan default (BADCRED)	.105	.258	- 4.25	<.0001
Mortgage payment to average rent	.726	.663	1.72	.0854
Liquid assets to mortgage payment	52.087	27.898	6.66	<.0001
Residual income to mortgage payment	3.096	2.709	2.88	.0040
Employment is positive factor	.311	.182	3.82	<.0001
Neighborhood vacancy > 5%	.281	.478	- 4.69	<.0001

Tests for Difference in Population Means for Approved Vs. Rejected Applicants

Variable	Approved Applicants	Rejected Applicants	t-Statistic	p-Value
Loan to value ratio	.803	.844	- 4.05	<.0001
Credit history is positive factor	.183	.052	5.15	<.0001
Previous loan default (BADCRED)	.107	.333	- 4.59	<.0001
Mortgage payment to average rent	.711	.800	- 1.22	.2224
Liquid assets to mortgage payment	50.343	33.079	4.03	<.0001
Residual income to mortgage Payment	3.053	2.974	.42	.6744
Employment is positive factor	.296	.204	.31	.7566
Neighborhood vacancy > 5%	.297	.417	- 2.29	.0220

Rejected applicants, on average, have significantly higher loan to value ratios and lower liquid asset to mortgage payment ratios than approved applicants. The average loan to value ratio is 80.3 percent for approved applicants compared to 84.4 percent for rejected applicants. Rejected applicants also have significantly worse credit histories than approved applicants (p-value < .0001). Approximately 18 percent of approved applicants have strong credit histories compared to only 5.2 percent of rejected applicants. One-third of rejected applicants have a history of loan defaults (BADCRED), and 10.7 percent of approved applicants have defaulted on loans (p-value < .0001). Previous research finds the applicant's credit history and liquid assets play an important role in loan approvals (Munnell *et. al.* (1996)). Previous default research finds that borrowers with higher loan to value ratios are more likely to default.

2. *Loan Default Data*

Descriptive statistics on the loan default dataset are shown in Table III. White borrowers, on average, have significantly lower loan to value ratios (p-value < .0001) and significantly higher liquid asset to mortgage payment ratios (p-value < .0001) than black borrowers. The average loan to value ratio is 100.2 percent for white borrowers compared to 101.2 percent for black borrowers. White borrowers in the sample are also more likely to have strong credit histories (p-value < .0001) and stable employment histories (p-value < .0001) than black borrowers. The borrower's credit history is considered a positive compensating factor for 22.8 percent of white borrowers and only 10.0 percent of black borrowers. Employment history is a positive compensating factor for 22.5 percent of white borrowers and 13.5 percent of black borrowers.

TABLE III

DESCRIPTIVE STATISTICS ON THE LOAN DEFAULT DATASET

Variable	Mean	Standard Deviation	Minimum	Maximum
All Borrowers (n=5000)				
Loan to value ratio	1.004	.054	.340	1.490
Credit history is positive factor	.202	.402	.000	1.000
Mortgage payment to average rent	1.622	.543	.438	4.855
Liquid assets to mortgage payment	9.652	11.261	.000	98.170
Residual income to mortgage payment	2.899	1.392	.050	13.870
Employment is positive factor	.207	.405	.000	1.000
Neighborhood vacancy > 5%	.622	.485	.000	1.000
White borrowers (n=4000)				
Loan to value ratio	1.002	.055	.340	1.480
Credit history is positive factor	.228	.419	.000	1.000
Mortgage payment to average rent	1.642	.543	.438	4.660
Liquid assets to mortgage payment	10.231	11.603	.000	98.170
Residual income to mortgage payment	2.878	1.394	.050	13.870
Employment is positive factor	.225	.418	.000	1.000
Neighborhood vacancy > 5%	.617	.486	.000	1.000
Black borrowers (n=1000)				
Loan to value ratio	1.012	.052	.670	1.490
Credit history is positive factor	.100	.300	.000	1.000
Liquid assets to mortgage payment	7.338	9.430	.000	90.890
Residual income to mortgage payment	2.981	1.379	.830	12.480
Employment is positive factor	.135	.342	.000	1.000
Neighborhood vacancy > 5%	.640	.480	.000	1.000

TABLE III (continued)

DESCRIPTIVE STATISTICS ON THE LOAN DEFAULT DATASET

Variable	Mean	Standard Deviation	Minimum	Maximum
Non-Defaulted Borrowers (n=2500)				
Loan to value ratio	.998	.062	.340	1.490
Credit history is positive factor	.397	.489	.000	1.000
Mortgage payment to average rent	1.674	.555	.508	4.855
Liquid assets to mortgage payment	11.556	12.181	.000	98.170
Residual income to mortgage payment	2.864	1.426	.050	13.870
Employment is positive factor	.397	.489	.000	1.000
Neighborhood vacancy > 5%	.599	.490	.000	1.000
Defaulted Borrowers (n=2500)				
Loan to value ratio	1.011	.044	.590	1.490
Credit history is positive factor	.008	.087	.000	1.000
Mortgage payment to average rent	1.570	.525	.438	4.724
Liquid assets to mortgage payment	7.749	9.902	.000	97.640
Residual income to mortgage Payment	2.934	1.356	.330	12.830
Employment is positive factor	.018	.132	.000	1.000
Neighborhood vacancy > 5%	.644	.479	.000	1.000

TABLE III (continued)

DESCRIPTIVE STATISTICS ON THE LOAN DEFAULT DATASET

Tests for Difference in Population Means for White Vs. Black Borrowers

Variable	White Borrowers	Black Borrowers	t-Statistic	p-Value
Loan to value ratio	1.002	1.012	- 5.38	<.0001
Credit history is positive factor	.228	.100	11.06	<.0001
Mortgage payment to average rent	1.642	1.542	5.28	<.0001
Liquid assets to mortgage payment	10.231	7.338	8.26	<.0001
Residual income to mortgage payment	2.878	2.981	- 2.11	.0348
Employment is positive factor	.225	.135	7.10	<.0001
Neighborhood vacancy > 5%	.617	.640	- 1.35	.1770

Tests for Difference in Population Means for Non-Defaulted Vs. Defaulted Borrowers

Variable	Non-Defaulted Borrowers	Defaulted Borrowers	t-Statistic	p-Value
Loan to value ratio	.998	1.011	- 8.55	<.0001
Credit history is positive factor	.397	.008	39.16	<.0001
Mortgage payment to average rent	1.674	1.570	6.81	<.0001
Liquid assets to mortgage payment	11.556	7.749	12.13	<.0001
Residual income to mortgage payment	2.864	2.934	- 1.78	.0750
Employment is positive factor	.397	.018	37.41	<.0001
Neighborhood vacancy > 5%	.599	.644	- 3.28	.0010

Defaulted borrowers, on average, have significantly higher loan to value ratios and lower liquid asset to mortgage payment ratios than borrowers that did not default. The average loan to value ratio is 101.1 percent for defaulted borrowers compared to 99.8 percent for borrowers that did not default. Borrowers that defaulted were less likely to have their previous credit history used as a positive compensating factor in the loan approval decision than borrowers that did not default. Only 0.8 percent of defaulted borrowers had credit histories that were considered positive compensating factors when the loan was originated. Almost 40 percent of borrowers that did not default had credit histories that were considered positive compensating factors when the loan was originated. There is also a significant difference in the employment history variable for defaulted and non-defaulted borrowers. Employment history was a positive compensating factor for only 1.8 percent of defaulted borrowers compared to 39.7 percent of borrowers that did not default. Previous research finds high loan to value ratios are positively related to loan defaults.

E. Statistical Methodology

1. Hypothesis One

Hypothesis one uses the owner wealth maximizing model and the loan default dataset to test whether the borrowing population is homogeneous or heterogeneous with respect to race. The following logit regression model is used on the loan default data to predict the probability of default for each loan:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e, \tag{20}$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan defaulted and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the loan amount to the lower of the purchase price or appraised value,

(2) a dummy variable that equals 1 if the applicant's credit history was a positive compensating factor in the loan decision and 0 otherwise,

(3) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(4) the ratio of the borrower's liquid assets to the mortgage payment,

(5) the ratio of the borrower's residual income to the mortgage payment,

(6) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise, and

(7) a dummy variable that equals 1 if the neighborhood vacancy rate exceeds 5 percent and 0 otherwise.

The probability of default, P_i, for each observation, is used as the dependent variable in an ordinary least squares regression model to determine if the borrower's race is a significant explanatory factor on the probability of default. The OLS regression model is shown:

$$P(DEFAULT) = b_0 + b_1 RACE + e \tag{21}$$

where P(DEFAULT) for each observation is equal to P_i from equation (20) and RACE equals 1 if the borrower is black and zero if the borrower is white.

If the coefficient on RACE is positive and significant, the average credit quality of black borrowers is less than the average credit quality of white borrowers. This would be evidence in support of the alternative hypothesis that the distribution of credit quality is heterogeneous with respect to race. The following empirical hypothesis is used to test whether the borrowing population is homogeneous or heterogeneous with respect to race:

$$H_0: b_1 \leq 0,$$
$$H_A: b_1 > 0.$$

2. *Hypothesis Two*

Hypothesis two tests the effectiveness of the equal outcome model in reducing racial disparities in lending decisions. This hypothesis is tested using the loan application data. A logit regression model is used to calculate the probability of loan approval, P_i, for each loan application using the owner wealth maximizing model and the equal outcome model. The logit model specification for the owner wealth maximizing model is shown:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e, \tag{22}$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan was approved and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the loan amount to the lower of the purchase price or appraised value,

(2) a dummy variable that equals 1 if the applicant's credit history was a positive compensating factor in the loan decision and 0 otherwise,

(3) a dummy variable that equals 1 if the borrower had a history of loan defaults and 0 otherwise,

(4) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(5) the ratio of the borrower's liquid assets to the mortgage payment,

(6) the ratio of the borrower's residual income to the mortgage payment,

(7) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise, and

(8) a dummy variable that equals 1 if the neighborhood vacancy rate exceeds 5 percent and 0 otherwise.

The logit model specification for the equal outcome model is:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e, \tag{23}$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan was approved and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(2) the ratio of the borrower's residual income to the mortgage payment, and

(3) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise.

The probability of loan approval, P_i, is calculated for each observation and used as the dependent variable in an ordinary least squares regression model to determine if the applicant's race is a significant explanatory factor on the probability of loan approval. Two separate regressions are used; one using the probability of loan approval, P_i, for each observation calculated using the owner wealth maximizing model and the other using the probability of loan approval, P_i, for each observation calculated using the equal outcome model. The OLS regression models are shown:

$$\begin{aligned} P(APPROVE) &= b_0 + b_1 RACE + e, \quad and \\ P(APPROVE) &= b_0^* + b_1^* RACE + e^*, \end{aligned} \tag{24}$$

where P(APPROVE) for each observation is equal to P_i from equation 22 for the owner wealth maximizing model and from equation 23 for the equal outcome model,

b_0, b_1, and e are the intercept, coefficent on RACE, and error term, respectively, from the owner wealth maximizing model, and b_0^*, b_1^*, and e^* are the intercept, coefficient on RACE, and error term, respectively, from the equal outcome model.

If minority borrowers have, on average, lower predicted probabilities of loan approval, the coefficient on RACE will be significant and negative. If the equal outcome model reduces the racial disparity in lending decisions, the coefficient on RACE should be greater (less negative) for the equal outcome model than for the owner wealth maximizing model. The empirical hypothesis is shown:

$$H_O:\ b_1^* \leq b_1,$$
$$H_A:\ b_1^* > b_1.$$

The parameter estimates b_1^* and b_1 and their variances are used to calculate a t-statistic to test for differences in the value of b_1^* and b_1 (Levin and Rubin (1998)):

$$t = \frac{b_1 * - b_1}{S_{b_1 * - b_1}} \tag{25}$$

where b_1^* is the coefficient on RACE for the equal outcome model, b_1 is the coefficient on RACE for the owner wealth maximizing model, and the pooled standard error is:

$$S_{b_1 * - b_1} = \sqrt{s_{b_1 *}^2 + s_{b_1}^2} \tag{26}$$

where $s_{b_1 *}^2$ is the variance for b_1* and $s_{b_1}^2$ is the variance for b_1.

A positive and significant t-statistic indicates the equal outcome model results in less racial disparity in lending decisions than the owner wealth maximizing model. Therefore, if the t-statistic is positive and significant, the null hypothesis is rejected in favor of the alternative hypothesis.

3. *Hypothesis Three*

Hypothesis three tests for differences in default prediction performance between the owner wealth maximizing model and the equal outcome model. This hypothesis is tested using the loan default dataset. A logit regression model is used to calculate the probability of default for each loan using the owner wealth maximizing model and the equal outcome model. The logit model specification for the owner wealth maximizing model is:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e, \tag{27}$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan defaulted and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the loan amount to the lower of the purchase price or appraised value,

(2) a dummy variable that equals 1 if the applicant's credit history was a positive compensating factor in the loan decision and 0 otherwise,

(3) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(4) the ratio of the borrower's liquid assets to the mortgage payment,

(5) the ratio of the borrower's residual income to the mortgage payment,

(6) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise, and

(7) a dummy variable that equals 1 if the neighborhood vacancy rate exceeds 5 percent and 0 otherwise.

The logit model specification for the equal outcome model is:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e, \tag{28}$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan defaulted and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(2) the ratio of the borrower's residual income to the mortgage payment, and

(3) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise.

The probability of default, P_i, for each observation and model are saved and recorded. The squared differences in the actual value of the dependent variable, DEFAULT, and the probability of default, P(DEFAULT), are used as the dependent variable in the following ordinary least squares regression model:[63]

$$\{DEFAULT - P(DEFAULT)\}^2 = b_0 + b_1 MODEL + e \tag{29}$$

where P(DEFAULT) for each observation is equal to P_i from equation (27) for the owner wealth maximizing model and from equation (28) for the equal outcome model, and MODEL equals 1 if the probability of default, P(DEFAULT), is from the owner wealth maximizing model and zero if the probability of default is from the equal outcome model. If the owner wealth maximizing model reduces the error in measuring the probability of default, the coefficient on MODEL should be less than zero. The empirical hypothesis used to test hypothesis three is shown:

$$H_O\text{: } b_1 \geq 0,$$
$$H_A\text{: } b_1 < 0.$$

[63] The prediction errors are squared so that large negative prediction errors and large positive prediction errors will not offset each other.

4. *Determination of the Costs of Using the Equal Outcome Model*

The equal outcome model should provide a benefit in terms of reducing the racial disparity in lending decisions. If the distribution of credit quality is heterogeneous with respect to race, the cost of using the equal outcome model is that lenders will be less able to distinguish good loans from bad loans than if they used the owner wealth maximizing model. This is the focus of hypothesis three. In addition to testing hypothesis three as described earlier, additional analysis is conducted to determine the costs of using the equal outcome model.

The probability of default, P_i, for each observation and model are used to determine if the loan is predicted to default or not default. Observations with a probability of default, P_i, greater than .5 are predicted to default, while observations with a probability of default less than .5 are predicted not to default.[64] The predicted outcomes are compared to the actual outcomes. The proportion of loans correctly classified is recorded for both empirical models. The following t-test is conducted to determine if the owner wealth maximizing model predicts loan defaults more accurately than the equal outcome model:

$$t = \frac{p_{OWMM} - p_{EOM}}{S_{p_{OWMM} - p_{EOM}}} \tag{30}$$

where p_{OWMM} is the correct prediction rate for the owner wealth maximizing model, p_{EOM}

[64] Aldrich and Nelson (1984) recommend this as one method for testing the goodness of fit of the model. Based on the sample proportions of 2500 defaulted loans and 2500 non-defaulted loans, the average probability of default should be .5. This method is presented in Gujarati (1995).

is the correct prediction rate for the equal outcome model, and

$$S_{P_{OWMM}-P_{EM}} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_{OWMM}} + \frac{\bar{p}(1-\bar{p})}{n_{EOM}}} \tag{31}$$

where n_{OWMM} is the sample size used with the owner wealth maximizing model, n_{EOM} is the sample size used with the equal outcome model, and

$$\bar{p} = \frac{n_{OWMM}\,p_{OWMM} + n_{EOM}\,p_{EOM}}{n_{OWMM} + n_{EOM}}. \tag{32}$$

Another measure is calculated to determine how much better the owner wealth maximizing model is than the equal outcome model in classifying good and bad loans. The reduction-in-error index (Klecka (1980), Huberty (1984) and Wilson and Sharda (1994) has been used to compare the performance of various models to pure chance models. With a slight modification, the reduction-in-error index can be used to compare the performance of the owner wealth maximizing model to the equal outcome model in reducing prediction errors. The calculation of the modified reduction-in-error statistic, I, is shown:

$$I = \frac{p_{OWMM} - p_{EOM}}{1 - p_{EOM}} \tag{33}$$

The reduction-in-error index, I, multiplied by 100 is the percentage fewer prediction errors using the owner wealth maximizing model instead of the equal outcome model.

The analysis of the cost of using the equal outcome model also includes a comparison of the dollar amount of loan defaults missed by the owner wealth maximizing model and by the equal outcome model. This is a simple comparison and does not involve statistical testing.

4. *Comparison of the Empirical Models to a Pure Chance Model in Loan Default Prediction*

The predictive validity of each model's classification versus a pure chance model is measured using the proportional chance criterion (Huberty (1984) and Wilson and Sharda (1994)). Each loan is classified as a predicted default or non-default as described in the previous section. The correct prediction rates are recorded and compared to the correct prediction rate that could be expected by pure chance. The proportional chance criterion (PCC) is based on the principle that one can achieve a correct prediction rate equal to the proportions in the sample. For example, if the sample consists of 10 percent defaulted loans and 90 percent non-defaulted loans one can achieve a 90 percent correct prediction rate by predicting none of the loans will default. The proportional chance criterion test statistic is distributed standard normal and is calculated as:

$$PCC = \frac{O - E\sqrt{N}}{\sqrt{E(N-E)}} \tag{34}$$

where O represents the total number of correct predictions, E is the total correct predictions obtainable by chance, and N is the total number of observations.

The reduction-in-error index is also used to compare the performance of the two empirical models to a pure chance model. The reduction-in-error index is computed as shown:

$$I = \frac{p_{MODEL} - p_{CHANCE}}{1 - p_{CHANCE}} \tag{35}$$

where p_{MODEL} is the correct prediction rate for the owner wealth maximizing model or the equal outcome model and p_{CHANCE} is the correct prediction rate obtainable by pure chance.

CHAPTER VI.
EMPIRICAL RESULTS

A. *Diagnostic Tests on the Regression Models*

Diagnostic tests include an evaluation of the effect of outlier observations, tests for multicollinearity and heteroscedasticity. The diagnostic tests are performed on the owner wealth maximizing model and the equal outcome model using both the loan application and loan default datasets.

The analysis of outlier observations on the model indicates the presence of several outlier observations in both datasets. The outlier observations do not appear to be the result of errors in the data. Therefore, they are retained in the sample.

Testing for multicollinearity focuses on the correlation coefficients for the independent variables. The correlation coefficients for the independent variables are shown in Table IV for both datasets. The only correlation coefficient above .5 is the correlation coefficient for credit history is a positive factor with employment is a positive factor in the loan default dataset. The majority of the correlation coefficients are below .10.

Gujarati (1995) outlines the consequences of multicollinearity. Models that have a high degree of multicollinearity may exhibit the following symptoms: (1) precise estimation may be difficult due to large variances and covariances, (2) due to large variances, the null hypothesis is more likely to be accepted, (3) individual t-ratios tend to be statistically insignificant, although the R^2 may be very high, and (4) the estimators and their standard errors may be sensitive to small changes in the data. The models used in

TABLE IV

CORRELATION COEFFICIENTS FOR THE INDEPENDENT VARIABLES

Loan Application Dataset:

	Loan to value ratio	Credit history is positive factor	Previous loan default
Loan to value ratio	1.000	- 0.203	0.053
Credit history is positive factor	- 0.023	1.000	- 0.174
Previous loan default	0.053	- 0.174	1.000
Mortgage payment to average rent	- 0.044	0.176	- 0.028
Liquid assets to mortgage payment	- 0.327	0.079	- 0.037
Residual income to mortgage payment	- 0.140	0.072	0.022
Employment is positive factor	0.050	- 0.084	- 0.057
Neighborhood vacancy rate (>5%)	0.111	- 0.047	0.054

	Mortgage payment to average rent	Liquid assets to mortgage payment
Loan to value ratio	- 0.044	- 0.327
Credit history is positive factor	0.178	0.079
Previous loan default	- 0.028	- 0.037
Mortgage payment to average rent	1.000	- 0.018
Liquid assets to mortgage payment	- 0.018	1.000
Residual income to mortgage payment	- 0.018	0.425
Employment is positive factor	- 0.093	- 0.035
Neighborhood vacancy rate (>5%)	- 0.043	- 0.018

	Residual income to mortgage payment	Employment is positive factor
Loan to value ratio	- 0.140	0.050
Credit history is positive factor	0.072	- 0.084
Previous loan default	0.022	- 0.057
Mortgage payment to average rent	- 0.018	- 0.093
Liquid assets to mortgage payment	0.425	- 0.035
Residual income to mortgage payment	1.000	- 0.011
Employment is positive factor	- 0.011	1.000
Neighborhood vacancy rate (>5%)	- 0.001	0.030

TABLE IV (continued)

CORRELATION COEFFICIENTS FOR THE INDEPENDENT VARIABLES

Loan Default Dataset:

	Loan to value ratio	Credit history is positive factor	Mortgage payment to average rent
Loan to value ratio	1.000	- 0.074	- 0.128
Credit history is positive factor	-0.074	1.000	0.023
Mortgage payment to average rent	-0.128	0.023	1.000
Liquid assets to mortgage payment	-0.372	0.110	0.008
Residual income to mortgage payment	0.102	0.011	- 0.238
Employment is positive factor	-0.052	0.703	0.033
Neighborhood vacancy rate (>5%)	0.013	- 0.044	0.103

	Liquid assets to mortgage payment	Residual income to mortgage payment
Loan to value ratio	- 0.372	0.102
Credit history is positive factor	0.110	0.011
Mortgage payment to average rent	0.008	- 0.238
Liquid assets to mortgage payment	1.000	0.036
Residual income to mortgage payment	0.036	1.000
Employment is positive factor	0.075	- 0.006
Neighborhood vacancy rate (>5%)	0.022	0.005

	Employment is positive factor	Neighborhood Vacancy rate (<5%)
Loan to value ratio	- 0.052	0.013
Credit history is positive factor	0.703	- 0.043
Employment is positive factor	1.000	- 0.017
Mortgage payment to average rent	0.033	0.103
Liquid assets to mortgage payment	0.075	0.022
Residual income to mortgage payment	- 0.006	0.005
Neighborhood vacancy rate (>5%)	- 0.017	1.000

this research do not exhibit any of these symptoms. Therefore, it appears that multicollinearity is not a serious problem.

Testing for heteroscedasticity in logit models is not well developed. This research uses the Breusch-Pagan-Godfrey test (Gujarati (1995)). The residuals are used to calculate the squared residual, e_i^2, and the variance of the residual:

$$\sigma_i^2 = \frac{\sum e_i^2}{n}$$

Another variable, p, is constructed as the ratio between the squared error and the variance of the residual:

$$p = \frac{e_i^2}{\sigma_{e_i}^2}$$

The variable, p, is used as the dependent variable in an OLS regression model where the independent variables are the same as the independent variables in the original regression model. The output from the regression on p is used to calculate the test statistic, ½(Model sum of squares). The test statistic is distributed chi-square with degrees of freedom equal to the number of independent variables in the model minus one.

The owner wealth maximizing model has heteroscedastic error terms using both datasets. For the loan application data, the model sum of squares is 33.96 and the Breusch-Pagan-Godfrey test statistic is 16.98 (significant at the .05 level). For the loan default data, the model sum of squares is 1238 and the test statistic is 619 (significant at the .0001 level).

The equal outcome model does not exhibit heteroscedastic error terms using the loan application data, but heteroscedasticity is present using the loan default data. For the loan application data, the model sum of squares is 1.5 and the Breusch-Pagan-Godfrey test statistic is .75 (not significant at the .10 level). For the loan default data, the model sum of squares is 1568 and the test statistic is 784 (significant at the .0001 level).

Several data transformations were attempted to remove or reduce the degree of heteroscedasticity. Data transformations attempted include (1) square root transformation (dividing through by the square root of the independent variable most likely to be causing the heteroscedasticity), and (2) inverse transformation. Both of the transformations increased the degree of heteroscedasticity. Therefore, the original model is used.

B. *Model Fit*

One standard measure used to evaluate model fit in regression models is the coefficient of multiple determination, R^2. Gujarati (1995) and Aldrich and Nelson (1984) argue that R^2 is of little practical use in models with qualitative dependent variables. An alternative test of model fit is the likelihood ratio test (Gujarati (1995)). The likelihood ratio test involves maximizing the likelihood function without any parameter restrictions, and then with restrictions that the coefficients in the model are all equal to zero. The test statistic, λ, is computed as shown:

$$\lambda = 2(\text{ULLF} - \text{RLLF}),$$

where ULLF is the log of the unrestricted likelihood function and RLLF is the log of the restricted likelihood function. The test statistic is distributed chi-square with degrees of freedom equal to the number of restrictions in the restricted model. This chi-square test statistic is computed for the owner wealth maximizing model and the equal outcome model for both datasets to evaluate model fit.

1. *Results of the Empirical Loan Default Models*

a. Owner Wealth Maximizing Model

The results of the owner wealth maximizing model on the loan default data are shown in Table V. The chi-square test statistic for the overall model fit is significant at the .0001 level.

The independent variables in the owner wealth maximizing model all have the expected sign except for the ratio of the monthly mortgage payment to the average neighborhood rent. The theoretical model indicates higher values of this ratio should be positively related to loan default, but the empirical results indicate higher values are negatively related to loan default. The ratio of the monthly mortgage payment to the average neighborhood rent represents the gross monthly mortgage payment, not the net payment. The net payment is the gross payment minus the tax benefits derived from making the payments. The net payment could not be computed for individual borrowers due to a lack of information on the tax brackets of individual borrowers. However, the tax benefit increases as the monthly mortgage payment increases. Therefore, the discrepancy in the expected sign and the observed sign on the ratio of the monthly mortgage payment to the average neighborhood rent may be the result of tax benefits associated with homeownership.

TABLE V

LOGIT REGRESSION RESULTS ON THE LOAN DEFAULT DATA

Owner Wealth Maximizing Model:

	Parameter Estimate	Standard Error	Chi-Square Test Statistic	P-Value
Intercept	- 1.311	0.800	2.69	.1013
Loan to value ratio	2.645	0.768	11.85	.0006
Credit history is positive factor	- 3.287	0.242	184.42	<.0001
Mortgage payment to average rent	- 0.370	0.064	33.17	<.0001
Liquid assets to mortgage payment	- 0.027	0.004	55.15	<.0001
Residual income to mortgage payment	- 0.001	0.026	0.01	.9704
Employment is positive factor	- 2.344	0.170	190.48	<.0001
Neighborhood vacancy rate (>5%)	0.217	0.071	9.42	.0021

The Chi-Square test statistic for the overall model fit is 1883 with 7 degrees of freedom ($p<.0001$).

Equal Outcome Model:

	Parameter Estimate	Standard Error	Chi-Square Test Statistic	P-Value
Intercept	1.059	0.138	59.05	<.0001
Residual income to mortgage payment	0.009	0.024	0.14	.7130
Mortgage payment to average rent	- 0.368	0.061	36.73	<.0001
Employment is positive factor	3.609	0.158	523.65	<.0001

The Chi-Square test statistic for the overall model fit is 1341.5 with 2 degrees of freedom ($p<.0001$).

All of the variables in the owner wealth maximizing model are significant at the .01 level except the ratio of residual income to the monthly mortgage payment. The p-value for the ratio of residual income to the monthly mortgage payment is .9704.

b. Equal Outcome Model

The results of the equal outcome model on the loan default data are shown in Table V. The chi-square test statistic for the overall model fit is significant at the .0001 level. The ratio of the mortgage payment to average neighborhood rent and the dummy variable that reflects employment history as a positive compensating factor are significant at the .0001 level with the expected signs. The ratio of residual income to the mortgage payment is not statistically significant.

2. *Results of the Empirical Loan Application Models*

a. Owner Wealth Maximizing Model

The results of the owner wealth maximizing model on the loan application data are shown in Table VI. The chi-square test statistic for overall model fit is significant at the .0001 level.

Only four of the empirical variables in the owner wealth maximizing model are significant at the .05 level. The four variables that are significant all have the expected sign. The variables that are significant at the .05 level are (1) the loan to value ratio, (2)

TABLE VI

LOGIT REGRESSION RESULTS ON THE LOAN APPLICATION DATA

Owner Wealth Maximizing Model:

	Parameter Estimate	Standard Error	Chi-Square Test Statistic	P-Value
Intercept	4.865	1.032	22.21	<.0001
Loan to value ratio	- 2.289	1.130	4.10	.0429
Credit history is positive factor	1.084	0.482	5.07	.0244
Previous loan default	- 1.255	0.244	26.46	<.0001
Mortgage payment to average rent	- 0.378	0.161	5.52	.0188
Liquid assets to mortgage payment	0.003	0.003	1.00	.3160
Residual income to mortgage payment	- 0.018	0.066	0.07	.7889
Employment is positive factor	0.086	0.244	0.13	.7234
Neighborhood vacancy rate (>5%)	- 0.402	0.225	3.20	.0735

The Chi-Square test statistic for the overall model fit is 56.1 with 7 degrees of freedom (p<.0001).

Equal Outcome Model:

	Parameter Estimate	Standard Error	Chi-Square Test Statistic	P-Value
Intercept	2.561	0.264	94.29	<.0001
Residual income to mortgage payment	0.026	0.063	0.16	.6851
Mortgage payment to average rent	- 0.219	0.155	1.98	.1593
Employment is positive factor	0.062	0.237	0.07	.7939

The Chi-Square test statistic for the overall model fit is 2.07 with 2 degrees of freedom (p=.5587).

the credit history is a positive factor dummy variable, (3) the previous loan default dummy variable, and (4) the ratio of the mortgage payment to average rent.

b. Equal Outcome Model

The results of the equal outcome model on the loan application data are shown in Table VI. The chi-square test statistic for overall model fit is insignificant. None of the independent variables in the model are significant at the .10 level.

C. Results of Hypothesis Tests

1. Hypothesis One: The Test for Heterogeneous Credit Quality in the Borrowing Population

Hypothesis one uses the owner wealth maximizing model and the loan default dataset to test whether the borrowing population is homogeneous or heterogeneous with respect to race. The following logit regression model is used on the loan default data to predict the probability of default for each loan:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e,$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan defaulted and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the loan amount to the lower of the purchase price or appraised value,

(2) a dummy variable that equals 1 if the applicant's credit history was a positive compensating factor in the loan decision and 0 otherwise,

(3) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(4) the ratio of the borrower's liquid assets to the mortgage payment,

(5) the ratio of the borrower's residual income to the mortgage payment,

(6) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise, and

(7) a dummy variable that equals 1 if the neighborhood vacancy rate exceeds 5 percent and 0 otherwise.

The probability of default, P_i, for each observation, is used as the dependent variable in an ordinary least squares regression model to determine if the borrower's race is a significant explanatory factor on the probability of default. The OLS regression model is shown:

$$P(DEFAULT) = b_0 + b_1 RACE + e$$

where P(DEFAULT) for each observation is equal to P_i from the previous equation and RACE equals 1 if the borrower is black and zero if the borrower is white.

If the coefficient on RACE is positive and significant, the average credit quality of black borrowers is less than the average credit quality of white borrowers. This would be evidence in support of the alternative hypothesis that the distribution of credit quality is heterogeneous with respect to race. The following empirical hypothesis is used to test whether the borrowing population is homogeneous or heterogeneous with respect to race:

$$H_O:\ b_1 \le 0,$$
$$H_A:\ b_1 > 0.$$

The test results for hypothesis one are shown in Table VII. The parameter estimate on RACE is .1061 and is significant at the .0001 level. The null hypothesis is rejected. The results support the alternative hypothesis that white borrowers, on average, have a higher distribution of credit quality than black borrowers. This indicates that the equal outcome model should reduce racial disparities in lending and that equal outcome lending should have a cost.

2. *Hypothesis Two: Racial Disparities in Lending Decisions*

Hypothesis two tests the effectiveness of the equal outcome model in reducing racial disparities in lending decisions. This hypothesis is tested using the loan application data. A logit regression model is used to calculate the probability of loan approval, P_i, for each loan application using the owner wealth maximizing model and the equal outcome model. The logit model specification for the owner wealth maximizing model is shown:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e,$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan was approved and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the loan amount to the lower of the purchase price or appraised value,

TABLE VII

RESULTS OF HYPOTHESIS TESTS

Hypothesis One: $P(DEFAULT) = b_0 + b_1 RACE + e$

	Parameter Estimate	Standard Error	t-Statistic	P-Value
Intercept	0.4788	0.0044	108.8	<.0001
Race	0.1061	0.0097	10.9	<.0001

Hypothesis Two: $P(APPROVE) = b_0 + b_1 RACE + e$

Owner Wealth Maximizing Model:

	Parameter Estimate	Standard Error	t-Statistic	P-Value
Intercept	0.9296	0.0018	514.9	<.0001
Race	- 0.0479	0.0051	- 9.4	<.0001

Equal Outcome Model:

	Parameter Estimate	Standard Error	t-Statistic	P-Value
Intercept	0.9235	0.0003	2659.4	<.0001
Race	0.0002	0.0010	0.2	.4207

$$t = \frac{.0002 - (.0479)}{.0052} = 9.25 \quad (p < .0001)$$

Hypothesis Three: $\{DEFAULT - P(DEFAULT)\}^2 = b_0 + b_1 MODEL + e$

	Parameter Estimate	Standard Error	t-Statistic	P-Value
Intercept	0.1933	0.0024	80.7	<.0001
Model	-0.0234	0.0034	- 6.9	<.0001

(2) a dummy variable that equals 1 if the applicant's credit history was a positive compensating factor in the loan decision and 0 otherwise,

(3) a dummy variable that equals 1 if the borrower had a history of loan defaults and 0 otherwise,

(4) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(5) the ratio of the borrower's liquid assets to the mortgage payment,

(6) the ratio of the borrower's residual income to the mortgage payment,

(7) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise, and

(8) a dummy variable that equals 1 if the neighborhood vacancy rate exceeds 5 percent and 0 otherwise.

The logit model specification for the equal outcome model is:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e,$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan was approved and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(2) the ratio of the borrower's residual income to the mortgage payment, and

(3) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise.

The probability of loan approval, P_i, is calculated for each observation and used as the dependent variable in an ordinary least squares regression model to determine if the applicant's race is a significant explanatory factor on the probability of loan approval. Two separate regressions are used; one using the probability of loan approval, P_i, for each observation calculated using the owner wealth maximizing model and the other using the probability of loan approval, P_i, for each observation calculated using the equal outcome model. The OLS regression models are shown:

$$P(APPROVE) = b_0 + b_1 RACE + e, \quad and$$
$$P(APPROVE) = b_0^* + b_1^* RACE + e^*,$$

where P(APPROVE) for each observation is equal to P_i from the owner wealth maximizing model or the equal outcome model, b_0, b_1, and e are the intercept, coefficent on RACE, and error term, respectively, from the owner wealth maximizing model, and b_0^*, b_1^*, and e^* are the intercept, coefficient on RACE, and error term, respectively, from the equal outcome model.

If minority borrowers have, on average, lower probabilities of loan approval, the coefficient on RACE will be significant and negative. If the equal outcome model reduces the racial disparity in lending decisions, the coefficient on RACE should be greater (less negative) for the equal outcome model than for the owner wealth maximizing model. The empirical hypothesis is shown:

$$H_O\text{: } b_1^* \leq b_1,$$
$$H_A\text{: } b_1^* > b_1.$$

The parameter estimates b_1^* and b_1 and their variances are used to calculate a t-statistic to test for differences in the value of b_1^* and b_1 (Levin and Rubin (1998)):

$$t = \frac{b_1 * - b_1}{S_{b_1 * - b_1}}$$

where b_1^* is the coefficient on RACE for the equal outcome model, b_1 is the coefficient on RACE for the owner wealth maximizing model, and the pooled standard error is:

$$S_{b_1 * - b_1} = \sqrt{s_{b_1 *}^2 + s_{b_1}^2}$$

where $s_{b_1 *}^2$ is the variance for b_1* and $s_{b_1}^2$ is the variance for b_1.

A positive and significant t-statistic indicates the equal outcome model results in less racial disparity in lending decisions than the owner wealth maximizing model. Therefore, if the t-statistic is positive and significant, the null hypothesis is rejected in favor of the alternative hypothesis.

The results of the OLS regression model used to test hypothesis two and the t-test to test if the coefficients on RACE are the same for both models are shown in Table VII. The parameter estimate on RACE is negative and significant for the owner wealth maximizing model. The RACE coefficient is not statistically significant for the equal outcome model. This indicates the equal outcome model does not produce significantly different probabilities of loan approval for white and black applicants. The t-test statistic indicates the parameter estimate on RACE is much greater (less negative) for the equal outcome model than for the owner wealth maximizing model. The t-statistic of 9.25 is significant at the .0001 level. Therefore, the null hypothesis is rejected. The evidence supports the alternative hypothesis that the equal outcome model results in less racial disparity than the owner wealth maximizing model.

3. *Hypothesis Three: Performance of the Models in Loan Default Prediction*

Hypothesis three tests for differences in default prediction performance between the owner wealth maximizing model and the equal outcome model. This hypothesis is tested using the loan default dataset. A logit regression model is used to calculate the probability of default for each loan using the owner wealth maximizing model and the equal outcome model. The logit model specification for the owner wealth maximizing model is:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e,$$

where,

$P_i = \text{PROB}(y_i = 1 \mid x_i)$,

y_i equals 1 if the loan defaulted and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the loan amount to the lower of the purchase price or appraised value,

(2) a dummy variable that equals 1 if the applicant's credit history was a positive compensating factor in the loan decision and 0 otherwise,

(3) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(4) the ratio of the borrower's liquid assets to the mortgage payment,

(5) the ratio of the borrower's residual income to the mortgage payment,

(6) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise, and

(7) a dummy variable that equals 1 if the neighborhood vacancy rate exceeds 5 percent and 0 otherwise.

The logit model specification for the equal outcome model is:

$$\log\left(\frac{P_i}{1-P_i}\right) = b_0 + b_i x_i + e,$$

where,

P_i = PROB(y_i = 1 | x_i),

y_i equals 1 if the loan defaulted and 0 otherwise,

and x_i is a vector of the following explanatory variables:

(1) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood,

(2) the ratio of the borrower's residual income to the mortgage payment, and

(3) a dummy variable that equals 1 if the borrower's employment history was a positive compensating factor in the loan decision and 0 otherwise.

The probability of default, P_i, for each observation and model are saved and recorded. The squared differences in the actual value of the dependent variable, DEFAULT, and the probability of default, P(DEFAULT), are used as the dependent variable in the following ordinary least squares regression model:

$$\{DEFAULT - P(DEFAULT)\}^2 = b_0 + b_1 MODEL + e$$

where P(DEFAULT) for each observation is equal to P_i calculated from the logit regression equations presented earlier, and MODEL equals 1 if the probability of default, P(DEFAULT), is from the owner wealth maximizing model and zero if the probability of default is from the equal outcome model. If the owner wealth maximizing model reduces the error in measuring the probability of default, the coefficient on MODEL should be less than zero. The empirical hypothesis used to test hypothesis three is shown:

$$H_O:\ b_1 \geq 0,$$
$$H_A:\ b_1 < 0.$$

The results of the OLS regression model used to test hypothesis three are shown in Table VII. The parameter estimate on MODEL is -.0234 and is significant at the .0001 level. This indicates the squared prediction errors are significantly lower for the owner wealth maximizing model than for the equal outcome model. Therefore, the null hypothesis is rejected. The evidence supports the alternative hypothesis that the owner wealth maximizing model predicts loan defaults more accurately than the equal outcome model.

D. *The Cost of Using the Equal Outcome Model*

The findings for hypothesis three indicate that the equal outcome model has larger errors in predicting loan default performance than the owner wealth maximizing model. Additional tests are conducted to determine if the owner wealth maximizing model results in fewer prediction errors when used to forecast whether or not the borrower will default. Based on the findings in hypothesis three, it is expected that the owner wealth maximizing model will have fewer prediction errors than the equal outcome model.

The probability of default, P_i, for each observation and model are used to determine if the loan is predicted to default or not default. Observations with a probability of default, P_i, greater than .5 are predicted to default, while observations with a probability of default less than .5 are predicted not to default. The predicted outcomes are compared to the actual outcomes. The proportion of loans correctly classified is recorded for both empirical models. The following t-test is conducted to determine if the owner wealth maximizing model predicts loan defaults more accurately than the equal outcome model:

$$t = \frac{p_{OWMM} - p_{EOM}}{S_{p_{OWMM} - p_{EOM}}}$$

where p_{OWMM} is the correct prediction rate for the owner wealth maximizing model, p_{EOM} is the correct prediction rate for the equal outcome model, and

$$S_{p_{OWMM} - p_{EM}} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_{OWMM}} + \frac{\bar{p}(1-\bar{p})}{n_{EOM}}}$$

where n_{OWMM} is the sample size used with the owner wealth maximizing model, n_{EOM} is the sample size used with the equal outcome model, and

$$\bar{p} = \frac{n_{OWMM} p_{OWMM} + n_{EOM} p_{EOM}}{n_{OWMM} + n_{EOM}}.$$

Results of the t-test for differences in the default prediction accuracy of the owner wealth maximizing model and the equal outcome model are shown in Table VIII. The t-statistic is 5.16 and is significant at the .0001 level. This indicates the owner wealth maximizing model outperforms the equal outcome model in distinguishing good loans from bad loans.

Another measure is calculated to determine how much better the owner wealth maximizing model is than the equal outcome model in classifying good and bad loans.

TABLE VIII

TESTS COMPARING THE PERFORMANCE OF THE TWO MODELS AND A PURE CHANCE MODEL IN LOAN DEFAULT PREDICTION

<u>The Cost of Using the Equal outcome model:</u>

Correct prediction rate for the Owner Wealth Maximizing Model = .736
Correct prediction rate for the Equal outcome model = .689

$$t = \frac{.736 - .689}{.0091} = 5.16 \quad (p < .0001)$$

Reduction-in-error index (I):

$$I = \frac{.736 - .689}{1 - .689} = .151$$

<u>Comparison of the Empirical Loan Default Models to a Pure Chance Model:</u>

Owner Wealth Maximizing Model:

Proportional Chance Criterion (PCC):

$$PCC = \frac{3680 - 2500\sqrt{2500}}{\sqrt{2500(5000 - 2500)}} = 33.4 \quad (p < .0001)$$

Reduction-in-error index (I):

$$I = \frac{.736 - .5}{1 - .5} = .472$$

Equal outcome model:

Proportional Chance Criterion (PCC):

$$PCC = \frac{3443 - 2500\sqrt{5000}}{\sqrt{2500(5000 - 2500)}} = 26.7 \quad (p < .0001)$$

Reduction-in-error index (I):

$$I = \frac{.689 - 5}{1 - .5} = .378$$

The reduction-in-error index (Klecka (1980), Huberty (1984) and Wilson and Sharda (1994) has been used to compare the performance of various models to pure chance models. With a slight modification, the reduction-in-error index can be used to compare the performance of the owner wealth maximizing model to the equal outcome model in reducing prediction errors. The calculation of the modified reduction-in-error statistic, I, is shown:

$$I = \frac{p_{OWMM} - p_{EOM}}{1 - p_{EOM}}$$

The reduction-in-error index, I, multiplied by 100 is the percentage fewer prediction errors using the owner wealth maximizing model instead of the equal outcome model.

As shown in Table VIII, using the owner wealth maximizing model results in 15.1 percent fewer prediction errors than using the equal outcome model. This provides additional support for the conclusion that the owner wealth maximizing model performs better in predicting loan defaults than the equal outcome model.

The analysis of the cost of using the equal outcome model also includes a comparison of the dollar amount of loan defaults missed by the owner wealth maximizing model and by the equal outcome model. This is a simple comparison and does not involve statistical testing.

The sample of 5,000 loans includes 2,500 that defaulted. The owner wealth maximizing model misclassified 1,320 loans, totaling $86,301,494. The equal outcome model misclassified 1,557 loans, totaling $99,507,963. For the data used in this research, the cost of using the equal outcome model is that 237 more loans, totaling $13,206,469, are misclassified than using the owner wealth maximizing model.

E. *Comparison of the Empirical Models to a Pure Chance Model in Loan Default Prediction*

The predictive validity of each model's classification versus a pure chance model is measured using the proportional chance criterion (Huberty (1984) and Wilson and Sharda (1994)). Each loan is classified as a predicted default or non-default as described earlier. The correct prediction rates are recorded and compared to the correct prediction rate that could be expected by pure chance. The proportional chance criterion (PCC) is based on the principle that one can achieve a correct prediction rate equal to the proportions in the sample. For example, if the sample consists of 10 percent defaulted loans and 90 percent non-defaulted loans one can achieve a 90 percent correct prediction rate by predicting none of the loans will default. The proportional chance criterion test statistic is distributed standard normal and is calculated as:

$$PCC = \frac{O - E\sqrt{N}}{\sqrt{E(N-E)}}$$

where O represents the total number of correct predictions, E is the total correct predictions obtainable by chance, and N is the total number of observations.

The reduction-in-error index is also used to compare the performance of the two empirical models to a pure chance model. The reduction-in-error index is computed as shown:

$$I = \frac{p_{MODEL} - p_{CHANCE}}{1 - p_{CHANCE}}$$

where p_{MODEL} is the correct prediction rate for the owner wealth maximizing model or the equal outcome model and p_{CHANCE} is the correct prediction rate obtainable by pure chance.

The results of the tests comparing the owner wealth maximizing model and the equal outcome model to a pure chance model are shown in Table VIII. The results indicate both the owner wealth maximizing model and the equal outcome model outperform a pure chance model.

The owner wealth maximizing model results in 47.2 percent fewer prediction errors than a pure chance model. The Proportional Chance Criterion test statistic of 33.4 is significant at the .0001 level.

The equal outcome model results in 37.8 percent fewer prediction errors than a pure chance model. The Proportional Chance Criterion test statistic of 26.7 is significant at the .0001 level.

CHAPTER VII.
CONCLUSIONS AND RECOMMENDATIONS

A. Conclusions

This research explores the possibility of developing an owner wealth maximizing lending model that will accurately assess credit risk without producing racial disparities in lending decisions. The empirical test results indicate that such a model probably does not exist because of the correlation of race with key economic characteristics of individual loan applicants. The findings indicate the distribution of credit quality is higher, on average, for white borrowers than for black borrowers.

Given that the ideal credit decision model doesn't exist, the research focuses on the tradeoffs between using a lending model that maximizes owner wealth and one that produces equal outcomes for white and black applicants. The empirical equal outcome model results in significantly less racial disparity in lending decisions than the owner wealth maximizing model.

While the equal outcome model reduces the racial disparity in lending decisions, there is a cost involved in using the equal outcome model. Policymakers should consider this cost as they move forward with performance-based measures in evaluating banks' lending efforts to minorities. If performance based measures lead to equal outcome lending, lenders may be forced to make loan decisions using inferior loan models. The equal outcome model in this research performed better than expected, but is still significantly worse than the owner wealth maximizing model in classifying good and bad loans.

Policymakers want to end racial discrimination in lending and that is an admirable and important goal. However, it will be difficult to remove statistical discrimination from the lending process because wealth maximizing lenders have an incentive to use any variable that they associate with creditworthiness in their lending model. Several of the variables associated with creditworthiness are correlated with race. If lenders include a variable in their model that is associated with creditworthiness that is correlated with race, the variable will have a disproportionate impact on minority applicants. It is difficult to disentangle discriminatory lending practices and owner wealth maximizing lending practices. One important first step in doing this is to verify that variables used in the loan decision model are based on some reasonable economic theory of the borrower's default decision and that empirically they are accurate measures of default risk.

B. Contributions of This Research

This investigation makes several contributions to the area of loan default and loan discrimination research. A key contribution of this research is the development of a new model of the borrower's default decision, which is incorporated into a loan decision model. The loan decision model was tested with loan default data to determine if the model can accurately assess credit risk before applying it to loan application data.

This research is the first attempt to validate a loan decision model by determining if the model accurately predicts loan defaults. Previous research identified variables the authors hypothesized to be important which were tested to determine if they explained loan approvals and denials. This investigation tests the loan decision model on loan default data first to verify that the model can accurately assess credit risk. It is important to determine that the model used in the loan decision process can accurately classify good and bad loans.

C. Limitations of This Research

The credit decision model developed in this research applies only to fixed-rate residential mortgage loans. It cannot be extended to adjustable-rate mortgage loans without significant revisions.

The findings for hypotheses one and three are limited to the U.S. lending market for single-family fixed rate residential mortgage loans. The findings for hypothesis two are limited to the Boston Metropolitan Statistical Area.

Data constraints made it difficult to proxy some of the factors in the theoretical model. The accuracy of the findings of this research are based on the assumption that the theoretical variables that could not be proxied would not drastically influence the empirical results.

D. Recommendations

Further research should be directed to extending the investigation of hypothesis two to other lending markets. A major problem with extending the investigation of hypothesis two to other lending markets is that the data needed to do so is generally not available.

Further research is also needed in the development of a credit decision model for adjustable-rate mortgage loans. Adjustable-rate mortgage loans play a prominent role in financing home mortgages. Developing a model for adjustable-rate single-family residential mortgage loans is difficult for several reasons. First, the monthly mortgage payment changes as interest rates change. The impact of this on the borrower's default decision may be dramatic. With an adjustable-rate mortgage, changes in interest rates result in changes in several of the empirical variables used in this research. For example, an increase in interest rates leads to higher monthly mortgage payments. Higher monthly mortgage payments result in changes in (1) the ratio of liquid assets to the mortgage payment , (2) the ratio of residual income to the mortgage payment, and (3) the ratio of the monthly mortgage payment to the average monthly rental rate for the neighborhood. Interest rate changes should play a significant role in the borrower's default decision on adjustable-rate mortgage loans, therefore any credit decision model of adjustable-rate mortgages must account for interest rate changes.

Finally, additional research is needed on the borrower specific characteristics that influence the probability of default. The loan underwriting guidelines used by major underwriters should be tested on loan default data to determine if all of the factors used in making loan decisions accurately measure default risk. If some of the factors are not good measures of default risk and the measures are correlated with race, the underwriting guidelines will have an unwarranted disproportionate impact on minority applicants.

REFERENCES

Aldrich, John H., and Forrest D. Nelson. 1984. "Linear Probability, Logit, and Probit Models." Sage University Working Paper, 1984.

Avery, Robert B., Patricia E. Beeson, and Mark S. Sniderman. 1992. "Cross-Lender Variation in Home Mortgage Lending." Working Paper No. 9219, Federal Reserve Bank of Cleveland (December).

__________. 1993a. "Lender Consistency in Housing Credit Markets." Working Paper No. 9309, Federal Reserve Bank of Cleveland (December).

__________. 1993b. "Accounting for Racial Differences in Housing Credit Markets." Working Paper No. 9310, Federal Reserve Bank of Cleveland (December).

Bacon, Kenneth H. 1993. "Clinton Expands Attack on Loan Bias as Data Shows a Continuing Problem." Wall Street Journal, (November 5), A-2.

Barefoot, JoAnn S. 1995. "What's a Fair Price?" ABA Banking Journal, (June): 28-33.

Bauer, Paul W., and Brian A. Cromwell. 1994. "A Monte Carlo Examination of Bias Tests in Mortgage Lending." Economic Review (Federal Reserve Bank of Cleveland), Vol. 30, No. 3 (Third Quarter): 27-40.

Becker, Gary S. 1971. The Economics of Discrimination, Second Edition. Chicago: University of Chicago Press.

Becker, Gary S. 1993. "The Evidence Against Banks Doesn't Prove Bias." Business Week, April 13.

Brealey, Richard A., and Stewart C. Myers. 1988. Principles of Corporate Finance. Third Edition. McGraw-Hill Book Company: New York, NY.

Brenner, Joel Glenn, and Liz Spayd. 1993. "A Pattern of Bias in Mortgage Loans." Washington Post, (June 6), A, 1:1.

Browne, Lynn E., and Geoffrey M.B. Tootell. 1995. "Mortgage Lending in Boston - A Response to the Critics." New England Economic Review, (Sept/Oct): 53-78.

Calem, Paul, and Michael Stutzer. 1995. "The Simple Analytics of Observed Discrimination in Credit Markets." Journal of Financial Intermediation, Vol. 4, No. 3 (July): 189-212.

Calomiris, Charles W., Charles M. Kahn, and Stanley D. Longhofer. 1994. "Housing-Finance Intervention and Private Incentive: Helping Minorities and the Poor." Journal of Money, Credit, and Banking, Vol. 26, No. 3 (August): 634-674.

Campbell, Tim S., and J. Kimball Dietrich. 1983. "The Determinants of Default on Conventional Residential Mortgages." Journal of Finance, Vol. 38, No. 5, (December): 1569-1581.

Canner, Glenn B., and Wayne Passmore. 1995. "Home Purchase Lending in Low-Income Neighborhoods and to Low-Income Borrowers." Federal Reserve Bulletin, (February): 71-103.

Canner, Glenn B., Wayne Passmore, and Dolores S. Smith. 1994. "Residential Lending to Low-Income and Minority Families: Evidence From the 1992 HMDA Data." Federal Reserve Bulletin, Vol. 80, (February): 79-108.

Canner, Glenn B., and Dolores S. Smith. 1991. "Home Mortgage Disclosure Act: Expanded Data on Residential Lending." Federal Reserve Bulletin, Vol. 77 (November): 859-881.

Canner, Glenn B., and Dolores S. Smith. 1992. "Expanded HMDA Data on Residential Lending: One Year Later. In HMDA Data and Mortgage Market Discrimination Research, ed. James H. Carr and Isaac F. Megbolugbe. Fannie Mae Research Roundtable Series, December 9, Washington, D.C.

Carr, James H., and Isaac F. Megbolugbe. 1993. "The Federal Reserve Bank of Boston Study on Mortgage Lending Revisited." Journal of Housing Research, 4(2): 277-313.

Cocheo, Steve. 1993. "Washington Post Slams Fair Lending Record." ABA Banking Journal, 58(8): 8-9.

England, Robert S. 1993. "Washington's New Numbers Game." Mortgage Banking, (September): 38-54.

Ferguson, Michael F., and Stephen R. Peters. 1995. "What Constitutes Evidence of Discrimination in Lending?" Journal of Finance, 50(2): 739-748.

Ferguson, Michael F., and Stephen R. Peters. 1997. "Cultural Affinity and Lending Discrimination: The Impact of Underwriting Errors and Credit Risk Distribution on Applicant Denial Rates." Journal of Financial Services Research, Vol. 11, No. 1 and 2 (February/April): 153-168.

Foster, C., and R. Van Order. 1984. "An Option Based Model of Mortgage Default." Housing Finance Review, 3(4): 351-372.

Galster, George C. 1993. "The Facts of Lending Discrimination Cannot Be Argued Away by Examining Default Rates." Housing Policy Debate, 4(1): 141-146.

Giliberto, S.M., and A.L. Houston. 1989. "Relocation Opportunities and Mortgage Default." Journal of the American Real Estate and Urban Economics Association, Vol. 17, No. 1: 55-69.

Gujarati, Damodar N. 1995. Basic Econometrics, Third Edition. McGraw-Hill, Inc.

Hendershott, Patric H., and William R. Schultz. 1993. "Equity and Nonequity Determinants of FHA Single-Family Mortgage Foreclosures in the 1980's." Journal of the American Real Estate and Urban Economics Association, Vol. 21, No. 4: 405-430.

Herzog, J.P., and J.S. Earley. 1970. Home Mortgage Delinquency and Foreclosure, New York: National Bureau of Economic Research.

Holmes, Andrew and Paul Horvitz. 1994. "Mortgage Redlining: Race, Risk and Demand." Journal of Finance, 49(1): 81-99.

Holmes, Andrew and Joe F. James. 1996. "Discrimination, Lending Practices, and Housing Values: Preliminary Evidence from the Houston Market." Journal of Real Estate Research, 11(1): 25-38.

Horne, David K. 1994. "Evaluating the Role of Race in Mortgage Lending." FDIC Banking Review, 7(1): 1-15.

Huberty, C.J. 1984. "Issues in the Use and Interpretation of Discriminant Analysis." Psychological Bulletin, 95: 156-171.

Hunter, William C., and Mary Beth Walker. 1995. "The Cultural Affinity Hypothesis And Mortgage Lending Decisions." Federal Reserve Bank of Chicago Working Paper No. 95-8 (July).

Jackson, Jerry R., and David L. Kaserman. 1980. "Default Risk on Home Mortgage Loans: A Test of Competing Hypotheses." Journal of Risk and Insurance, (December): 679-690.

Kau, James B., Donald C. Keenan, and Taewon Kim. 1994. "Default Probabilities for Mortgages." Journal of Urban Economics, Vol. 35: 278-296.

Klecka, W.R. 1980. Discriminant Analysis. Sage Publishing: Beverly Hills, CA.

LaCour-Little, Michael. 1996. "Application of Reverse Regression to Boston Federal Reserve Data Refutes Claims of Discrimination." Journal of Real Estate Research, Vol. 11, No. 1: 1-12.

Leahy, Peter J. 1985. "Are Racial Factors Important for the Allocation of Mortgage Money? A Quasi Experiment Approach to the Aspect of Discrimination." American Journal of Economics and Sociology, Vol. 44, No. 2: 185-199.

Lekkas, Vassilis, John M. Quigley and Robert Van Order. 1993. "Loan Loss Severity And Optimal Mortgage Default." Journal of the American Real Estate and Urban Economics Association, Vol. 21: 353-371.

Levin, Richard I., and David S. Rubin. 1998. Statistics for Management, Seventh Edition. Prentice Hall.

Liebowitz, S. 1993. "A Study That Deserves No Credit." Wall Street Journal, September 1, A14.

Longhofer, Stanley D. 1996. "Cultural Affinity and Mortgage Discrimination." Economic Review (Federal Reserve Bank of Cleveland), Vol. 32, No. 3 (Third Quarter): 12-24.

Munnell, Alicia H., Geoffrey M.B. Tootell. Lynn E. Browne, and James McEneaney. 1996. "Mortgage Lending in Boston: Interpreting HMDA Data." American Economic Review, 86(1): 25-53.

Nesiba, Reynold F. 1996. "Racial Discrimination in Residential Lending Markets: Why Empirical Researchers Always See It and Economic Theorists Never Do." Journal of Economic Issues, Vol. 30, No. 1 (March): 51-77.

Perle, Eugene D., Kathryn Lynch and Jeffrey Horner. 1993. "Model Specification and Local Mortgage Market Behavior." Journal of Housing Research, 4(2): 225-244.

Phillips-Patrick, Fred J., and Clifford V. Rossi. 1996. "Statistical Evidence of Mortgage Redlining? A Cautionary Tale." Journal of Real Estate Research, 11(1): 13-24.

Riddiough, Timothy J. 1991. "Equilibrium Mortgage Default Pricing with Non-Optimal Borrower Behavior." Ph.D. Dissertation. University of Wisconsin-Madison.

Schwartz, Eduardo S., and Walter N. Torous. 1993. "Mortgage Prepayment and Default Decisions: A Poisson Regression Approach." Journal of the American Real Estate and Urban Economics Association, Vol. 21, No. 4: 431-449.

Quercia, Roberto G., and Michael A. Stegman. 1992. "Residential Mortgage Default: A Review of the Literature." Journal of Housing Research, 3(2): 341-379.

Schill, Michael H., and Susan M. Wachter. 1993. "A Tale of Two Cities: Racial and Ethnic Geographic Disparities in Home Mortgage Lending in Boston and Philadelphia." Journal of Housing Research, 4(2): 245-275.

Shaffer, Sherrill. 1996. "Evidence of Discrimination in Lending: An Extension." Journal of Finance, 51(4): 1551-1554.

Stiglitz, J.E., and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." American Economic Review, (June): 393-410.

Tootell, Geoffrey M.B. 1993. "Defaults, Denials, and Discrimination in Mortgage Lending." New England Economic Review, (Sept/Oct): 45-51.

Tootell, Geoffrey M.B. 1996. "Redlining in Boston: Do Mortgage Lenders Discriminate Against Neighborhoods?" Quarterly Journal of Economics, (November): 1049-1079.

Vandell, Kerry D. 1978. "Default Risk Under Alternative Mortgage Instruments." Journal of Finance, 33(5): 1279-1298.

Vandell, Kerry D., and Thomas Thibodeau. 1985. "Estimation of Mortgage Defaults Using Disaggregate Loan History Data." Journal of the American Real Estate and Urban Economic Association, 13(3): 292-316.

von Furstenberg, George M. 1969. "Default Risk on FHA-Insured Home Mortgages as a Function of the Terms of Financing: A Quantitative Analysis." Journal of Finance, 24: 459-477.

Wilson, Rick L., and Ramesh Sharda. 1994. "Bankruptcy Prediction Using Neural Networks." Decision Support Systems, 11: 545-557.

Zandi, Mark. 1993. "Boston Fed's Bias Study Was Deeply Flawed." American Banker, August 19.

OKLAHOMA STATE UNIVERSITY
INSTITUTIONAL REVIEW BOARD
HUMAN SUBJECTS REVIEW

Date: 04-23-97 IRB#: BU-97-011

Proposal Title: A COST-BENEFIT ANALYSIS OF ECONOMICALLY OPTIMAL VERSUS RACE NEUTRAL RESIDENTIAL MORTAGE LENDING

Principal Investigator(s): W. Gary Simpson, Christopher L. Brown

Reviewed and Processed as: Exempt

Approval Status Recommended by Reviewer(s): Approved

ALL APPROVALS MAY BE SUBJECT TO REVIEW BY FULL INSTITUTIONAL REVIEW BOARD AT NEXT MEETING, AS WELL AS ARE SUBJECT TO MONITORING AT ANY TIME DURING THE APPROVAL PERIOD.
APPROVAL STATUS PERIOD VALID FOR DATA COLLECTION FOR A ONE CALENDAR YEAR PERIOD AFTER WHICH A CONTINUATION OR RENEWAL REQUEST IS REQUIRED TO BE SUBMITTED FOR BOARD APPROVAL.
ANY MODIFICATIONS TO APPROVED PROJECT MUST ALSO BE SUBMITTED FOR APPROVAL.

Comments, Modifications/Conditions for Approval or Disapproval are as follows:

Signature: John H. Wyckoff

Chair of Institutional Review Board

Date: April 24, 1997

cc: Christopher L. Brown

VITA

Christopher L. Brown

Candidate for the Degree of

Doctor of Philosophy

Dissertation: AN ANALYSIS OF OWNER WEATLH MAXIMIZING VERSUS EQUAL OUTCOME RESIDENTIAL MORTGAGE LENDING

Major Field: Business Administration

Biographical:

Education: Graduated from North Little Rock Ole Main High School, North Little Rock, Arkansas in May, 1979; received Bachelor of Business Administration degree in Marketing and a Master of Business Administration degree from the University of Central Arkansas, Conway, Arkansas, in May, 1983 and May, 1986, respectively. Completed the degree requirements for the Doctor of Philosophy degree with a major in Finance at Oklahoma State University in May, 1998.

Experience: Employed as a savings and loan examiner with the Office of Thrift Supervision from April, 1986 to August, 1991. Employed as a Graduate Teaching Assistant at Oklahoma State University from August, 1991 to May, 1995 and employed as a Visiting Assistant Professor of Finance at Southeastern Louisiana University from August, 1995 to the present.

Professional Memberships: Financial Management Association.